WP/10/230

# IMF Working Paper

# Monetary Policy Matters: New Evidence Based on a New Shock Measure

*S. Mahdi Barakchian and Christopher Crowe*

INTERNATIONAL MONETARY FUND

 WP/10/230

Research Department

**Monetary Policy Matters: New Evidence Based on a New Shock Measure**

**Prepared by S. Mahdi Barakchian and Christopher Crowe[1]**

Authorized for distribution by Atish Ghosh

October 2010

**Abstract**

Conventional VAR and non-VAR methods of identifying the effects of monetary policy shocks on the economy have found a negative output response to monetary tightening using U.S. data over the 1960s-1990s. However, we show that these methods fail to find this contractionary effect when the sample is restricted to the period since the 1980s, apparently due to changes in the policymaking environment that reduce their effectiveness. Identifying policy shocks using Fed Funds futures data, we recover the contractionary effect of monetary tightening on output and find that almost half of output variation over the period appears due to policy shocks.

JEL Classification Numbers:C32; E52; G13

Keywords: Monetary policy; VAR estimation; Fed Funds Futures; FOMC

Author's E-Mail Address:CCrowe@imf.org; barakchian@sharif.ir



[1] [1] Barakchian: Assistant Professor, Sharif University of Technology, Tehran, Iran. Crowe (corresponding author): International Monetary Fund, Research Department. ccrowe@imf.org. +1 202 623-5303. This project was commenced while Barakchian was a summer intern in the IMF's Research Department. The authors would like to thank, subject to the usual caveats, Olivier Blanchard and Larry Christiano for useful discussions and conference and seminar participants at the IMF, Bank of England, Banco de España and the Federal Reserve Bank of Chicago, Julian Di Giovanni, Ken Kuttner, Hashem Pesaran, Paolo Surico, Eric Swanson and—in particular—David Romer for useful comments on an earlier draft.

Contents

## I. INTRODUCTION

Identifying the impact of monetary policy on the economy is a central question in empirical macroeconomics. The key identification problem is simultaneity. Hence, the focus has been on the exogenous or 'shock' component of policy changes. For the U.S., a consensus has emerged on the qualitative effects of a monetary policy shock. Christiano, and others, (1999) summarize this consensus as follows:

> After a contractionary monetary policy shock, short term interest rates rise, aggregate output, employment, profits and various monetary aggregates fall, the aggregate price level responds very slowly, and various measures of wages fall, albeit by very modest amounts. In addition, there is agreement that monetary policy shocks account for only a very modest percentage of the volatility of aggregate output; they account for even less of the movements in the aggregate price level.

These results are in line with the predictions of benchmark models used for policy analysis. The consensus holds across different means of identification, including recursive VARs (e.g. Christiano, and others, 1996) and non-recursive VARs (e.g. Sims and Zha, 2006) and non-VAR identification (e.g. Romer and Romer's narrative approach; 1989 and 2004).[2]

However, as we demonstrate, this consensus is sensitive to the period used for analysis. In particular, it is dependent on the inclusion of the 1970s and early 1980s, when shocks were large and the policymaking environment was very different from the one faced today. When one attempts to identify the effects of monetary policy shocks for the period since the 1980s using the same methodologies one obtains quite different results. Notably, contractionary monetary policy shocks appear to have a small positive effect on output.

We present some evidence on changes to the nature of U.S. monetary policy shocks that would cause conventional identification methods to give misleading results. First, we show that U.S. monetary policy has become more systematic, responding more to the variables in policymakers' information set, so that the signal/noise ratio of the shock component of policy actions has shrunk, making it harder to identify the effect of such shocks. Second, we show that policymaking has become more forward looking. Hence, VAR identification methods that ignore the role of forecasts in the policymaker's reaction function are misspecified.[3] Identification methods (such as Romer and Romer, 2004) that allow for forward-looking variables in the reaction function but do not allow for the apparent increase in their relative weight will tend to suffer from the same problem. For instance, a monetary contraction aimed at partially offsetting

[2]Contradicting this consensus, Uhlig (2005) imposes sign restrictions on the impulse responses of prices, nonborrowed reserves and the federal funds rate in response to a monetary policy shock and shows that the effect of a contractionary monetary policy shock on output is unclear. Similarly, Ozlale (2003) contradicts the consensus view that monetary policy shocks have little impact on output, finding that around 65 percent of output variation can be attributed to monetary policy shocks (more in line with our results).

[3]More generally, we present evidence of structural breaks in the monetary policy reaction function which suggests that all identification strategies that impose a time-invariant reaction function are misspecified.

an anticipated positive output shock will show up as a lagged positive output response to a monetary contraction if the forward-looking aspect of policymaking is not appropriately allowed for. This could explain the apparently perverse output response uncovered for the more recent period using conventional identification methods.

We turn to financial market data in an effort to uncover a measure of monetary policy shocks that is less subject to these criticisms. Following Kuttner (2001), Gürkaynak, Sack and Swanson (2005) and Piazzesi and Swanson (2008) we identify monetary policy shocks as the 'surprise' component of monetary policy actions, estimated using movements in Fed Funds Futures contract prices on the day of monetary policy announcements following FOMC meetings.

The Fed Funds Futures market has been in operation since late 1988. Contracts are available for several months into the future, so that information on the surprise component of the policy announcement can be obtained from the contract for the current calendar month as well as future months. One benefit of using a range of futures contracts—not simply that for the current month—is that the shock to the current month's futures rate can simply reflect resolved uncertainty about the timing of the policy change, rather than the overall direction of policy (Gürkaynak, 2005 and Bernanke and Kuttner, 2005). To efficiently capture the information contained across the maturity spectrum, we use factor analysis to uncover the common information from six monthly contracts: the current month and up to 5 months ahead. Similarly to Gürkaynak, and others (2005) who apply a factor model to a set of eurodollar and fed funds futures contracts with a maturity structure of up to one year ahead, we find that two factors are sufficient to summarize the information across the six contracts. Moreover, similarly to the literature on factor models of the yield curve (e.g. Piazzesi, 2010), the factors have a natural interpretation as level and slope, respectively. We use the former as our measure of the policy shock.

We enter this new shock measure in a simple monthly VAR, similarly to Romer and Romer (2004), estimated for 1988:12-2008:06.[4] We find that, with our new measure, a contractionary monetary policy shock has a statistically significant negative effect on output. While the effect is small in absolute terms, forecast error variance decomposition suggests that, in an era of low overall output volatility, our new policy shock measure can account for up to half of output volatility at a horizon of 3 years or more—around twice the proportion using existing shock measures. We find some evidence for a 'price puzzle': contractionary monetary policy also leads to a small, and borderline significant, increase in the general price level at a horizon of 1-3 years, although this is subsequently reversed. Efforts to eliminate the price puzzle by including a measure of commodity prices or inflation expectations in the VAR, following suggestions in the literature, are not successful.

[4]Because the Fed Funds futures market only started trading in October 1988, we are unable to derive our shock measure for the early portion of the great moderation. However, the results for the other identification strategies we follow in section II are broadly the same whether the estimation starts in 1982, 1984 or 1988. We end the sample at the end of the second quarter of 2008 since the intensification of the global financial crisis in the third quarter of 2008 (following the collapse of Lehman Brothers) likely represents a significant structural break. Our results are also robust to ending the sample at the end of 2007.

The principal benefit of using the surprise component of policy announcements as a proxy for the policy shock is that one eliminates all the predictable (public information) elements in the policy reaction function whose inclusion could bias our estimate of the impact of policy. Moreover, this method imposes no restrictions on the variables in the reaction function or its functional form. However, to the extent that the Fed has accurate private information about the future state of the economy, simultaneity bias could still be a problem. This is because the surprise component of policy announcements combines two separate pieces of 'news'. One is the policy shock. The other is news about the Fed's private information set. With the policy surprise used as a proxy for the shock, the estimated impact of the shock on output will tend to be biased if the Fed's private information is accurate, because the response of policy to the Fed's private forecasts of macro variables will be falsely interpreted as the response of the variables to policy. Hence, the policy surprise is a reasonable proxy for the policy shock—and will deliver unbiased estimates of the impact of monetary policy—only to the extent that the surprise is orthogonal to the policymaker's information set.

To assess this, we regress our shock measure on the Fed's private information set, using Romer and Romer's specification for the Fed's reaction function, and proxying for the Fed's private information using the difference between the Fed's private (Greenbook) forecast and publicly-available private sector forecasts (*Blue Chip* forecasts) for each variable. We find that the Fed's private information can explain less than 19 percent of the shock measure, while the joint null hypothesis of zero coefficients on all 17 variables in the private information set cannot be rejected (p-value .13). Hence, the surprise component of the policy announcement captured in our measure seems a good proxy for a monetary policy shock.[5]

Our methodology builds on the insights of an increasingly influential literature on identifying monetary policy shocks using financial market data. Rudebusch (1998) was an early paper advocating the use of Fed Funds futures data, while Kuttner's (2001) focus on one-day changes in futures prices, rather than the difference between the implied futures rate and the actual policy rate, allowed for sharper identification. Faust, and others (2004) propose a novel two-stage identification scheme in which the information available from the Fed Funds futures is used to partially identify a structural VAR. Gürkaynak, and others (2005) use a two factor model to combine information from futures contracts (both Fed Funds futures and Eurodollar futures) at different horizons and separately identify level and slope factors. Hamilton (2008) derives level, slope and curvature factors using three Fed Funds futures contracts, and estimates the impact of the different factors on housing market variables. Thapar (2008) uses 3 month Treasury Bill futures prices as a proxy for market expectations, in a novel identification method that combines

[5]However, looking at coefficients on individual variables, we find evidence that our shock measure reacts positively to the Fed's private forecasts of near-term economic developments, specifically current-quarter output and inflation. If, as a result, our 'shock' measure includes the Fed tightening in response to its private information on near-term output pressures, then the estimated effect of the policy tightening on output will be biased, to the extent that the Fed's forecasting advantage is real. Romer and Romer's (2000) results suggest that the Fed does indeed enjoy a forecasting advantage, and our analysis of the Fed and private sector forecasts supports this. However, this bias is likely to be positive, so that our estimated negative effect should be an *under-estimate* (in absolute terms). Since the Fed's forecasting advantage is found to be relatively slight, the bias will likely be small. These issues are discussed in more detail in section V.

these market-based forecasts with Greenbook forecasts of output and price variables.

While we are therefore not the first to consider these methods, we believe that our particular identification scheme offers some advantages, and is well-suited to address the research questions we are seeking to address. Because, in our case, the policy shock is identified outside the VAR, we are able to avoid some of the weaknesses of structural VAR estimation outlined in Section III (the lack of forward-looking variables, existence of structural breaks). By contrast, Faust and others (2004) use the structural VAR model to identify the monetary policy shock and to estimate the impulse responses of the macro variables to the policy shock, and as a result their method is subject to some of these criticisms. Like Kuttner (2001) and Hamilton (2008), but unlike Rudebusch (1998) and Thapar (2008), we focus on daily innovations in Fed Funds futures prices. Using daily data from policy announcement days helps to remove the impact of other news (such as economic data releases) and more cleanly identify the impact of exogenous policy shocks. Moreover, as Kuttner (2001) has argued, focusing on innovations to the futures price helps to strip out the impact of fluctuations in term and risk premia. Our focus on the information in futures contracts up to six months out contrasts with Gürkaynak and others (2005), who analyze contracts up to twelve months out, and Hamilton (2008), who analyzes the three nearest term futures contracts. While the choice of horizon is somewhat arbitrary and in our case is mainly dictated by the degree of liquidity in contracts at different maturities, we believe that 6 months is roughly the right horizon for policy considerations.[6] Finally, our approach, as well as being extremely intuitive, is somewhat easier to implement and to reproduce for other applications than that of Faust and others (2004) and Thapar (2008), and we hope that our estimated shock series will be widely used by other researchers.

In the next section, we briefly review the literature on identifying monetary policy shocks and their effects. We focus in particular on four identification schemes that have received significant attention: Christiano and others' recursive VAR identification (1996); Sims and Zha's (2006) non-recursive VAR; Bernanke and Mihov's (1998) over-identified VAR; and Romer and Romer's (2004) narrative identification. We contrast the baseline results in the original papers with results for the recent period (focusing on the post-1988 period to allow a comparison with our new measure). In section III, we analyze how the nature of monetary policy shocks has changed since the early 1980s, using Romer and Romer's (2004) specification of the Fed's reaction function and information set to show how policy has become both more deterministic in general and more forward-looking in particular. In section IV we discuss the Fed Funds Futures market and outline our new shock measure. In section V we use our new measure to estimate the effects of monetary policy shocks in the post-1988 period, discuss the results and outline some robustness checks. Section VI concludes.

---

[6]Our approach has some additional advantages: unlike Hamilton (2008), it extracts the underlying information from the futures contracts using an unrestrictive functional form, and, in extracting two factors from six contracts rather than three factors from three contracts, is less demanding on the data. By focusing solely on Fed Funds futures contracts, rather than combining these with futures based on longer-maturity money market rates as in Gürkaynak and others (2005), we avoid the additional complications created by the inclusion of policy and non-policy rates together. For instance, the emergence of a significant time-varying spread between the policy rate and money market rates, due to financial market stress, towards the end of our sample would create significant noise if innovations to money market futures were used to infer policy shocks.

## II. Conventional identification schemes

### A. Identifying Monetary Policy Shocks

Following Christiano and others (1999), we identify a monetary policy shock as the orthogonal disturbance term $s_t$ in an equation of the form:

$$S_t = f(\Omega_t) + s_t \tag{1}$$

where $S_t$ denotes the monetary stance (or more narrowly, the instrument of the monetary authority, e.g. the Fed Funds rate) and $f$ is a linear function relating $S_t$ to the policymaker's information set $\Omega_t$ [7]

We focus on this exogenous shock component in order to avoid the simultaneity bias that arises when elements of the Fed's information set $\Omega_t$ are also endogenous variables whose response to the policy stance we want to estimate. For instance, assume the following two equation system, where the policy stance responds to the central bank's estimate of output and output responds negatively to policy:

$$\begin{aligned} S_t &= \varphi E_{CB}[Y_t] + s_t \\ Y_t &= -\alpha S_t + u_t \end{aligned} \tag{2}$$

Assume that the central bank's forecast of the output shock $u_t$ (denoted $\widehat{u}_t$) has some informational content (i.e. $Cov(\widehat{u}_t, u_t) > 0$), but the central bank does not know the policy shock $s_t$ in advance.[8] Then the solution to this model is given by:

$$\begin{aligned} S_t &= \frac{\varphi}{1+\alpha\varphi}\widehat{u}_t + s_t \\ Y_t &= -\frac{\alpha\varphi}{1+\alpha\varphi}\widehat{u}_t - \alpha s_t + u_t \end{aligned} \tag{3}$$

Then regressing $Y_t$ on $S_t$ will give a biased estimate of $\alpha$, since $Cov(S_t, u_t) = \frac{\varphi}{1+\alpha\varphi} Cov(\widehat{u}_t, u_t) > 0$ (simultaneity bias). The bias will be positive (that is, if the true impact of monetary policy is contractionary, a smaller contractionary impact or a positive effect

---

[7]Hence equation (1) can be thought of as the monetary authorities' feedback rule or policy reaction function, although as Christiano and others (1999) highlight, there are pitfalls in identifying the coefficients in $f(\,)$.

[8]The policy shock $s_t$ is assumed orthogonal to the other disturbance terms: $Cov(u_t, s_t) = Cov(\widehat{u}_t, s_t) = 0$.

will be found in the data):

$$E\left[-\widehat{\alpha}\right] = -\alpha + \frac{\varphi}{1+\alpha\varphi} \frac{Cov\left(\widehat{u_t}, u_t\right)}{\left(\frac{\varphi}{1+\alpha\varphi}\right)^2 Var\left(\widehat{u_t}\right) + Var\left(s_t\right)} \tag{4}$$

The bias disappears if monetary policy does not in fact respond to the Fed's private information ($\varphi = 0$); if the Fed's private information is garbage ($Cov\left(\widehat{u_t}, u_t\right) = 0$); or the variance of the monetary policy shock explodes ($Var\left(s_t\right) \to \infty$). However, a regression of $Y_t$ on $s_t$, the monetary shock, will give an unbiased estimate since $Cov\left(s_t, u_t\right) = 0$ by definition.

Conventional VAR identification schemes identify monetary policy shocks by estimating a VAR with sufficient identifying assumptions to uncover the structural parameters and shocks. For instance, consider the following reduced form VAR model:

$$\mathbf{z}_t = B(L)\mathbf{z}_{t-1} + \mathbf{v}_t \tag{5}$$

where $\mathbf{z}_t$ is a $n \times 1$ vector of variables, $\mathbf{v}_t$ is a vector of reduced form errors and $\text{var}(\mathbf{v}_t) = \Sigma$. A corresponding structural form of this VAR can be written as:

$$A_0\mathbf{z}_t = A(L)\mathbf{z}_{t-1} + \boldsymbol{\varepsilon}_t \tag{6}$$

where $A_0$ is a matrix of contemporaneous coefficients, $\boldsymbol{\varepsilon}_t$ is a vector of structural shocks and $\text{var}(\boldsymbol{\varepsilon}_t) = \Psi$. Therefore, $B(L) = A_0^{-1}A(L)$, $\mathbf{v}_t = A_0^{-1}\boldsymbol{\varepsilon}_t$ and $\Sigma = A_0^{-1}\Psi A_0^{-1'}$.

Since $B(L)$ and $\Sigma$ are computed through estimation, in order to recover structural parameters, $A_0$, $A(L)$, and $\Psi$, we need to impose $n^2$ restrictions on the structural coefficients for exact identification. These restrictions are usually imposed on $A_0$ and/or $\Psi$.[9] For identifying monetary policy shocks the key issue is which variables enter contemporaneously in $\Omega_t$ .

---

[9]The most widely used identification scheme is the recursive approach (Sims, 1980). In this approach, $\Psi$ is diagonal—the shocks are treated as orthogonal—which gives $n(n-1)/2$ restrictions, $A_0$ is (lower) triangular, which gives $n(n-1)/2$ restrictions, and normalizing the diagonal elements of $A_0$ gives $n$ more restrictions. The triangular assumption on $A_0$ implies a Wold-causal ordering by which each variable in $\mathbf{z}_t$ is a function of the contemporaneous values of the variables above but not below it. This ordering is typically motivated by some theory about the relative timing of economic activities and decision making. For non-recursive VARs the difference is simply that while $n(n-1)/2$ restrictions are still imposed on the contemporaneous coefficients matrix, $A_0$ is not assumed to be lower triangular. That is, there is some simultaneity assumed in the contemporaneous relationships, motivated by economic theory. Another strand of the VAR literature identifies shocks via restrictions on long-run coefficients (see e.g. Blanchard and Quah, 1989).

If the monetary instrument $S_t$ is the $i^{th}$ element of $\mathbf{z}_t$, equation (1) can be proxied by the $i^{th}$ row of (SVAR)—with the $i^{th}$ element of $\boldsymbol{\varepsilon}_t$ providing an estimate of the monetary policy shock $s_t$.

### B. Results for four identification schemes: comparing the recent period with earlier results.

Here we outline results for four representative identification schemes, replicating the results for the original period under analysis in each case, and comparing with results for our baseline period (1988-2008). The four schemes we consider are Christiano, Eichenbaum and Evans' (1996) recursive VAR approach, Bernanke and Mihov's (1998) over-identified VAR, Sims and Zha's (2006) non-recursive VAR and Romer and Romer's (2004) narrative approach. The full details of these approaches and our efforts to replicate them are detailed in the appendix. This section provides a brief overview of the results.

Estimated over their original sample periods—from the 1960s to the mid-1990s—all four approaches suggest that monetary policy shocks have an effect in line with the conventional wisdom from DSGE models: a monetary contraction lowers output and other real indicators over the short to medium term, and has a more muted impact—generally negative—on inflation. However, estimating the models over the more recent period yields very different results. Most worryingly, monetary contractions are estimated to have a stimulative effect on output.

**Christiano, Eichenbaum and Evans (1996)** estimate a quarterly VAR with six variables and four lags over the period 1960Q1-1992Q4. Their results show that following a contractionary monetary shock, the federal funds rate rises and various measures of money fall. They also show that a contractionary shock is associated with a persistent decline in output. The price index responds slowly but eventually declines.[10] We replicate their results and report the impulse responses of output and price with two standard error bands after a contractionary monetary policy shock (Figure 1 panel a).[11] However, when we estimate the same model for the recent period (1988Q4-2007Q3), neither output nor prices show the expected response (Figure 1 panel b).[12] After a contractionary monetary policy shock output increases significantly, while prices show no significant increase or decrease.

**Bernanke and Mihov (1998)** develop a model in which the relationships among macroeconomic variables are left unrestricted while contemporaneous identification restrictions are imposed on monetary variables in order to model the Fed's operating procedure. We re-estimate their model

---

[10]This result is different from some other studies, for example Eichenbaum (1992) and Sims (1992), who find evidence for a price puzzle: i.e. a prolonged increase in the price index following a contractionary monetary policy shock. In order to avoid this puzzle, Christiano et. al. (1996), like Bernanke and Mihov (1998) and Sims and Zha (2006), assume that the monetary authority reacts to commodity prices in setting monetary policy. They show that when the commodity price index is excluded from the VAR, the price puzzle reemerges. Including a commodity price index for the recent period has no effect on the response of consumer prices to policy shocks (section V).

[11]In this paper the size of the monetary policy shock is always equal to one standard deviation and impulse responses are always reported with two standard error bands. Standard errors are obtained via multivariate normal parametric bootstrapping, based on 500 replications.

[12]We end our sample in 2007 Q3 because nonborrowed reserves (NBR) become negative during the fourth quarter of 2007. Our sample is also truncated (at 2007:11) for the Bernanke and Mihov estimation for the same reason.

for the original period (1965:01-1996:12), and also for our period of interest (1988:12-2007:11). Figure 2 (panel a) shows that in the original period the responses of output and prices are as expected, and very similar to Christiano and others': following a contractionary monetary shock output falls and prices fall with a delay but greater persistence.[13] However, when we estimate this model for the later period (panel b), again neither output nor prices show the expected response. Both output and prices increase significantly—immediately in the case of output, and over the medium term in the case of prices.[14]

**Sims and Zha (2006)** include a somewhat different list of variables from most other studies, including the producers' price index components for crude materials and intermediate materials and a measure of bankruptcies. We replicate their findings for their original sample (Figure 3, panel a).[15] After a contractionary shock all the price indices eventually fall and output declines, similarly to the results using Christiano and others' (1996) recursive identification scheme (the results are not significant due to the wide standard error bands obtained under the bootstrap algorithm). However, when we estimate the model for the 1988:Q4-2008:Q2 period (panel b), the impulse responses are very different. After a contractionary monetary policy shock, output increases significantly over the medium term.

**Romer and Romer (2004)** argue that these means of identifying monetary policy shocks are subject to two major deficiencies (a failure to control for anticipated monetary policy changes and for deviations between desired and actual changes due to endogenous movements in monetary instruments), and develop a narrative approach that seeks to overcome these problems. Romer and Romer estimate a monthly VAR with three variables: the log of industrial production, log PPI for finished goods and their measure of the monetary policy shock derived through their narrative method.[16] For their original sample (1969:01-1996:12) they find that a monetary policy shock has large, relatively rapid, and statistically significant effects on both output and inflation, and the effects of their new measure are substantially stronger and quicker than for conventional measures of monetary policy. Figure 4 (panel a) illustrates their findings. However, when we

[13]Bernanke and Mihov estimate different versions of the model, including four that are over-identified and one that is just-identified. We replicate the over-identified model (Federal Funds rate targeting model) since Bernanke and Mihov find that this performs best for the post-1988 period.

[14]Although we re-estimate the same VAR, i.e. a monthly VAR with 13 lags and six variables (output, domestic prices, commodity prices, the Federal Funds rate, total reserves and NBR), there are some minor differences between our VAR and Bernanke and Mihov's. They interpolate GDP and the GDP Deflator to convert a quarterly series to a monthly series, while we use monthly Industrial Production and CPI data instead. We also use a different commodity price index. We believe that these differences are minor, and comparing the impulse responses from the original period suggests that they have no significant effect on the results.

[15]Due to data constraints, we exclude their bankruptcy measure from the VAR. The impulse responses of our model estimated for the original period are almost identical to those in Sims and Zha (2006). In fact Sims and Zha mention that the measure of bankruptcy makes only a modest contribution to the results, while Christiano and others (1999) also re-estimate the Sims and Zha model excluding the bankruptcy measure. Having said this, our confidence intervals are somewhat wider than those reported by Sims and Zha: this is partly cosmetic (they report 68 percent, or approximately one standard error, CIs, whereas we report two standard error CIs); it may also reflect the exclusion of the bankruptcy measure in our estimates, or possibly differences in the bootstrap algorithms.

[16]Since the Federal Funds rate enters in levels in the VAR analysis, Romer and Romer cumulate the new shock measure to produce a comparable series. They also estimate single-equation specifications and find similar results (see Romer and Romer, 2004). The VAR includes 36 lags of the endogenous variables, a constant and a linear time trend.

estimate this model for the period 1988:12-2008:06, the impulse responses are different, especially for output (panel b).[17] After a contractionary monetary policy shock, the price level goes down, but the response is not as strong as for the original period. The output response is initially flat, but with a significant positive effect after around 2 years.

### C. Discussion

What can we take from these findings? The overall message is that the results using the existing identification strategies seem to be sensitive to the sample period. Of particular concern for current policymakers, the results for the most recent—and presumably most relevant—period appear to be out of line with the theoretical consensus, especially for output. However, we argue that there are good reasons to doubt the robustness of these empirical results. Several identification problems are likely to have become particularly acute for the recent period. In the following section, we provide some evidence for this.

## III. Evolution of Federal Reserve Policymaking and Policy Shocks

Each of the identification strategies outlined in the previous section estimates a version of the policy reaction function outlined in equation (1). Romer and Romer estimate it explicitly using elements of the Fed's private information set as proxies for $\Omega_t$, and identify the monetary policy shocks $s_t$ with the residuals. The structural VAR identification methods estimate the reaction function as the $i^{th}$ equation in the VAR system, where the elements of $\Omega_t$ depend on the assumptions made about the contemporaneous coefficients matrix $A_0$, and monetary policy shocks are identified as the $i^{th}$ element of the orthogonalized residuals matrix $\varepsilon_t$. In each case, a key assumption is that both the elements of $\Omega_t$ and the coefficients in $f()$ are correctly identified. We show, using Romer and Romer's specification for $f(\Omega_t)$ as a benchmark, that these assumptions are likely to be invalid.[18] We then discuss what implications these findings have for the conventional identification results presented in section II.

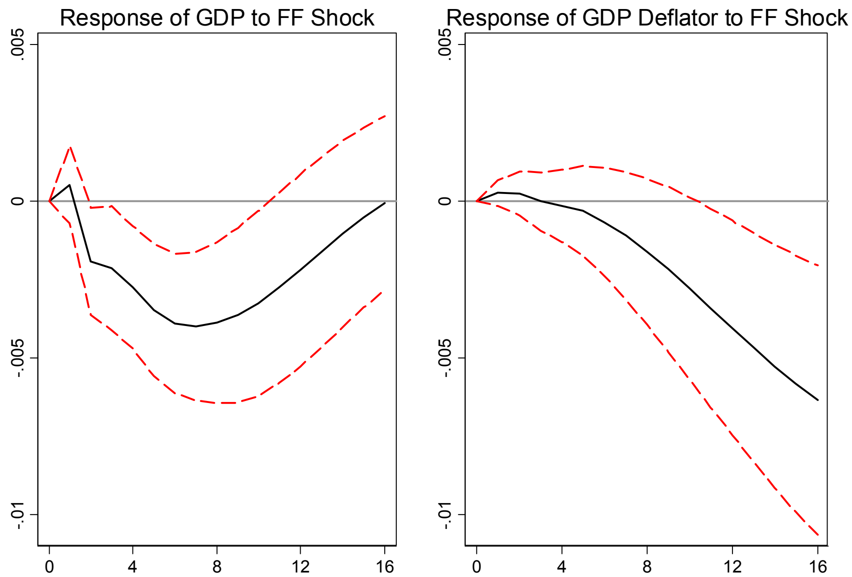
Romer and Romer's reaction function has the change in the desired Fed Funds target rate as the dependent variable. The right hand side variables include the level of the desired Fed Funds target going into the FOMC meeting in question, and 17 forecast variables taken from the Greenbook forecasts. The latter include the current quarter unemployment rate estimate, and Eight estimates/forecasts for real GDP growth and the change in the GDP deflator respectively.

[17] See the data Appendix for information on how the Romer and Romer index was extended to 2008.

[18] Of course this is not the only reaction function one could use. The reaction function literature is voluminous; see, for instance, Taylor, 1993; Orphanides, 2003; Clarida and others 1999, 2000. However, Romer and Romer's approach has received considerable attention in the literature, while the authors themselves show in a series of robustness checks that, from the point of view of the shocks series, different permutations of the rule yield similar results.

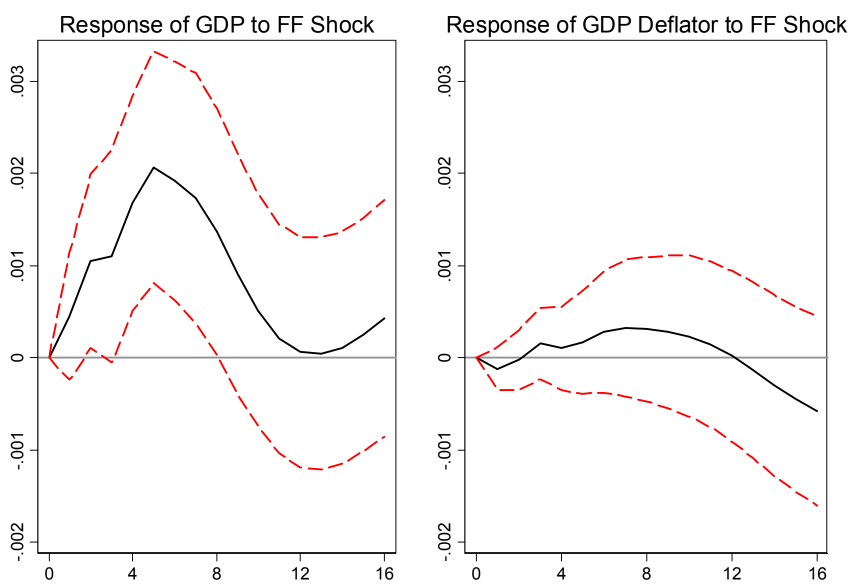
**Figure 1. Christiano and others**
**Panel a.**

Christiano, Eichenbaum and Evans. 1960Q1-1992Q4




**Panel b.**

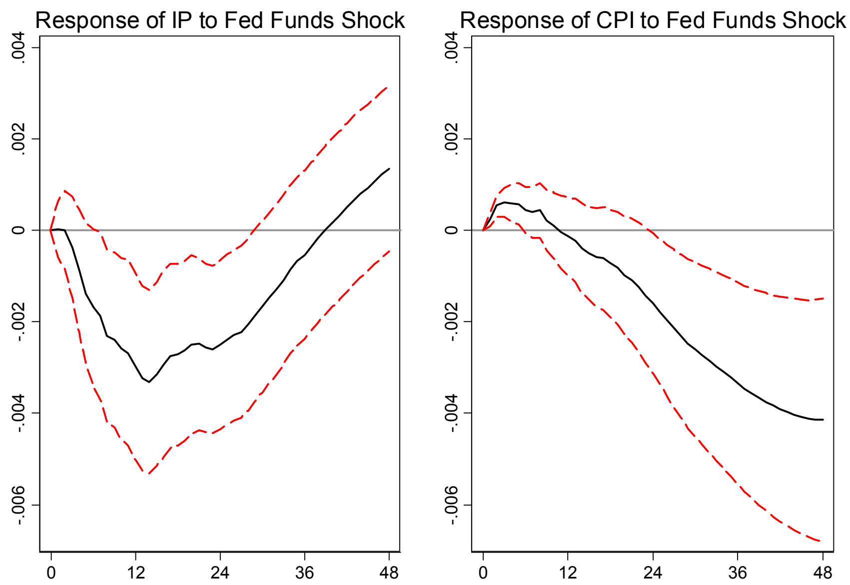
Christiano, Eichenbaum and Evans. 1988Q4-2007Q3




Structural VAR (quarterly data, 6 endogenous variables plus constant and linear time trend, 4 lags) as described in text.
Variables ordered as GDP, GDP deflator, commodity prices, non-borrowed reserves, Fed Funds rate, total reserves.
All variables except for the Fed Funds rate are in logs and seasonally adjusted. Graphs show response of GDP and GDP deflator to a one standard deviation positive shock to the Fed Funds rate.
Structural shocks obtained via Cholesky decomposition.
Two Standard Error bands produced by parametric bootstrapping (500 replications).

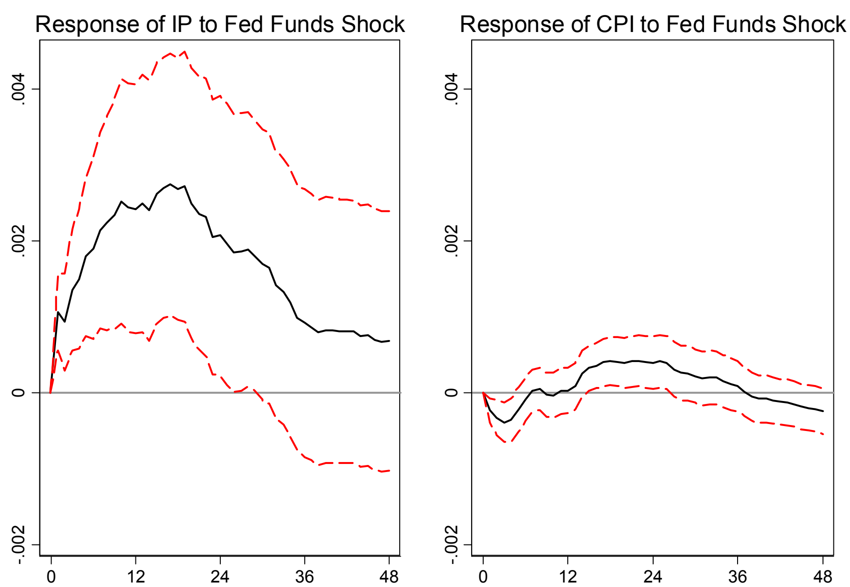
**Figure 2. Bernanke and Mihov**
**Panel a**

Bernanke and Mihov. 1965:01-1996:12




**Panel b**

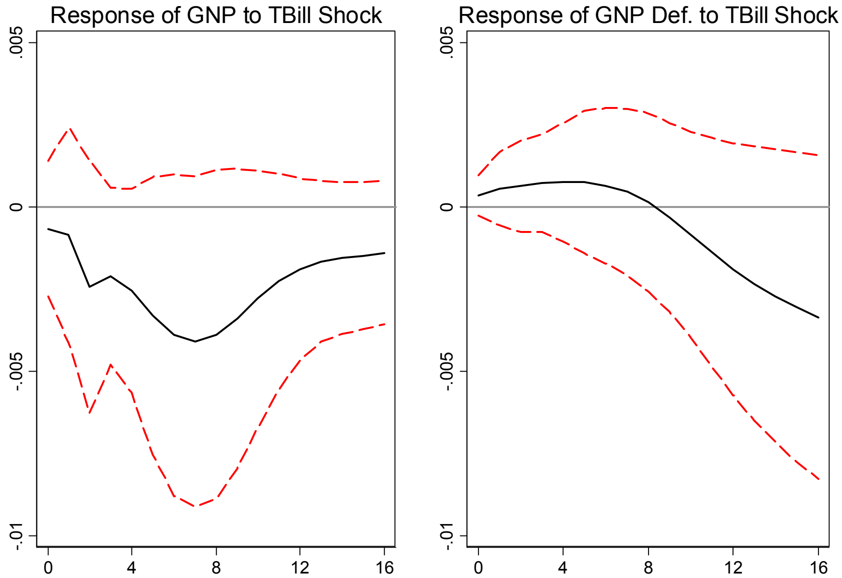
Bernanke and Mihov. 1988:12-2007:11




Structural VAR (monthly data, 6 endogenous variables plus constant and linear time trend, 13 lags) as described in text.
Variables include industrial production, consumer price index, commodity prices, Fed Funds rate, total reserves, non-borrowed reserves. The first 3 variables are in logs and seasonally adjusted. The last two variables are seasonally adjusted and normalized by dividing by the 36-month moving average of total reserves. Graphs show response of output and CPI to a one standard deviation positive shock to the Fed Funds rate. Structural Shocks obtained by imposing the structural decomposition discussed in the text (1 overidentifiying restriction)
Two Standard Error bands produced by parametric bootstrapping (500 replications).

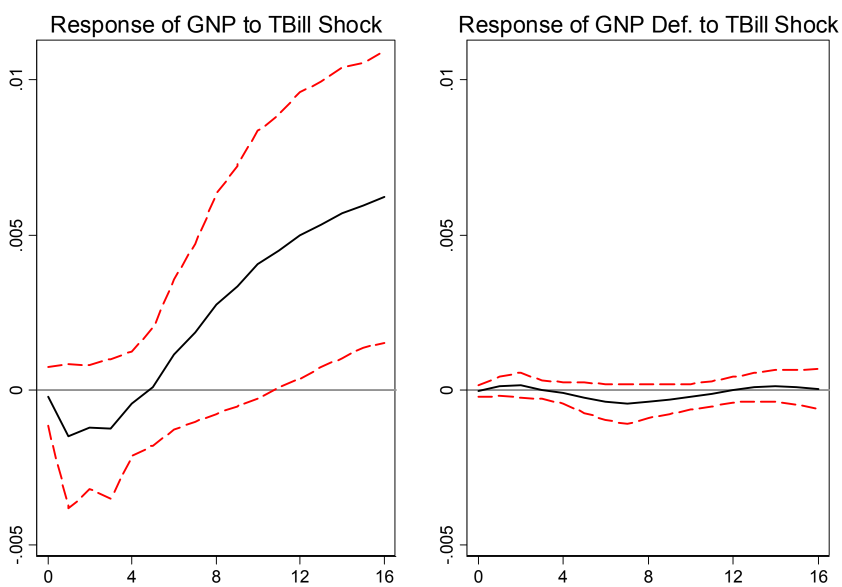
## Figure 3. Sims and Zha
## Panel a.

Sims and Zha. 1964:Q1-1994:Q4



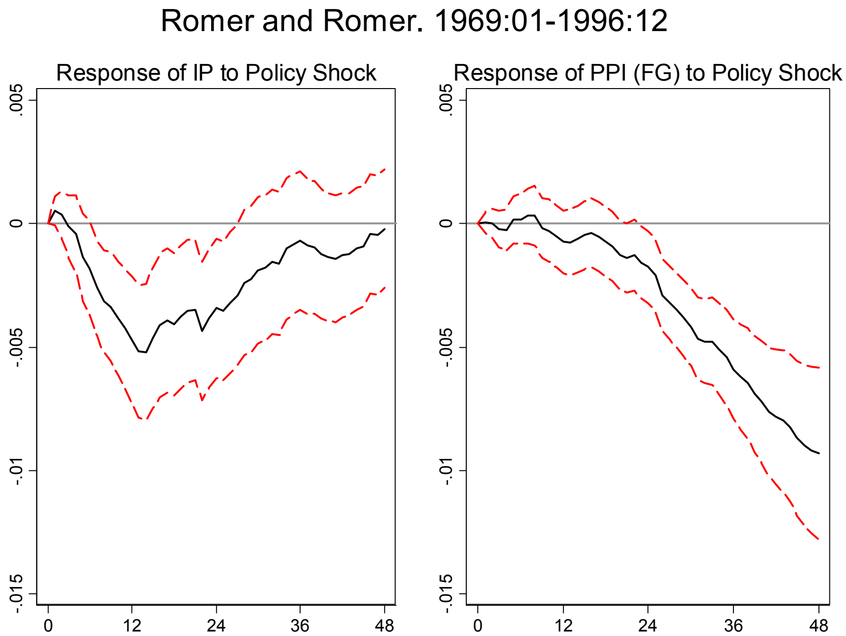

## Panel b.

Sims and Zha. 1988:Q4-2008:Q2

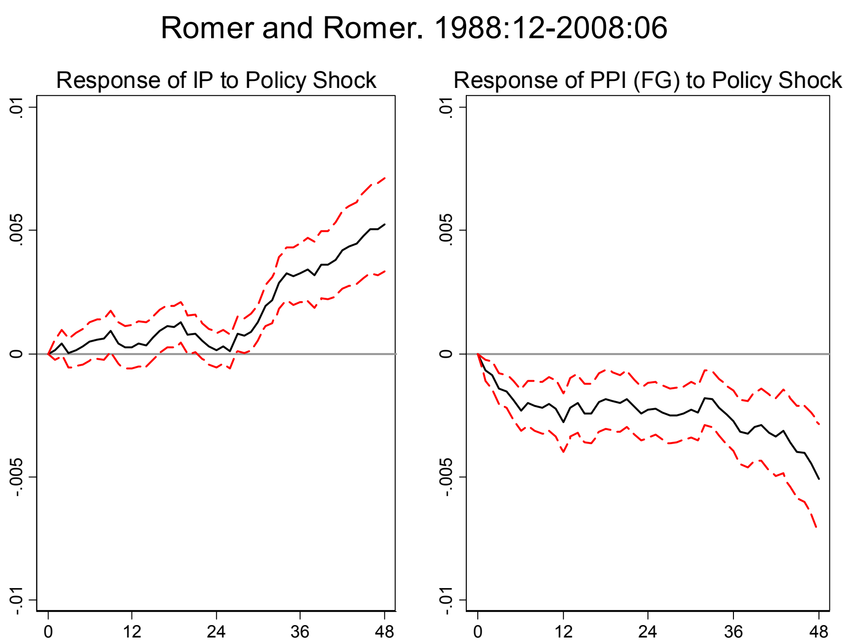



Structural VAR (Quarterly data, 7 endogenous variables plus constant and linear time trend, 4 lags) as described in text.
Variables include Crude Materials Prices, M2, T Bill Rate, Intermediate Materials Prices, GNP Deflator, Wages (private sector workers) and GNP. All variables except the T Bill Rate are in logs and seasonally adjusted. Graphs show response of GNP and GNP Deflator to a one standard deviation positive shock to the T Bill Rate.
Structural Shocks obtained by imposing the structural decomposition discussed in the text (2 overidentifying restrictions).
Two Standard Error bands produced by parametric bootstrapping (500 replications).

## Figure 4. Romer and Romer
## Panel a.




## Panel b.

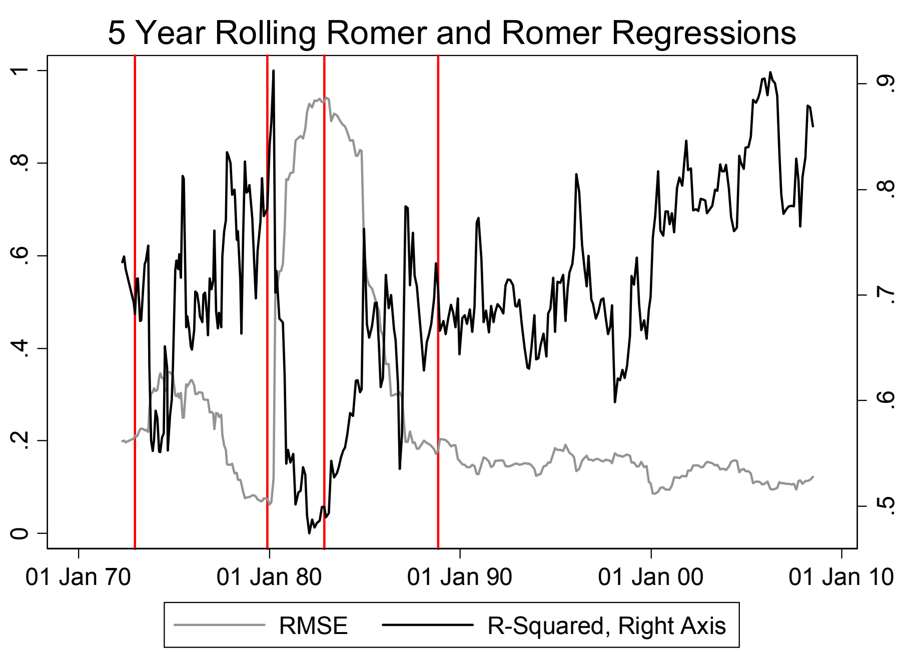



Structural VAR (Monthly data, 3 endogenous variables plus constant and linear time trend, 36 lags).
Variables ordered as industrial production, producer price index (finished goods), both seasonally adjusted and in logs, and Romer and Romer's shock measure, cumulated. Graphs show response of industrial production and PPI (finished goods) to a one standard deviation positive shock to the policy measure.
Structural shocks obtained via Cholesky decomposition.
Two Standard Error bands produced by parametric bootstrapping (500 replications).

In each case, these eight variables include estimates for the current and previous quarters, forecasts for the next two quarters following the meeting, and, for each of these four variables, the change in each estimate between the last meeting and the current one.

To give an overview of how Romer and Romer's policy regression performs over time, we analyze the variance of the estimated shock series and the regression's goodness of fit. Figure 5 plots the root mean squared error (RMSE) and $R^2$ from rolling regressions using the Romer and Romer specification estimated over rolling 40-meeting (approximately 5 year) windows (the date corresponds to the end of the relevant window). The RMSE—which gives the standard deviation of the estimated shock series for each 5 year window—peaks in the mid-1970s following the first oil shock, then spikes dramatically in the wake of the Volcker shock before declining more or less monotonically to the end of the sample. The share of the variation in policy rates explained by the deterministic part of the reaction function follows a mirror image, declining from around .75 in the late 1970s to below .5 in the wake of the Volcker shock, then increasing to .8-.9 in recent years.

**Figure 5. Results from 5 year Rolling Regressions**




Root mean squared error (RMSE) and $R^2$ from rolling 40-observation regressions of Romer and Romer's policy reaction function (observations organized by meeting, March 1969-June 2008). Vertical lines delimit the subsamples identified by Bagliano and Favero (1998).

This illustrates a general problem that will tend to have reduced the effectiveness of all identification strategies. With policymaking becoming more deterministic in recent years, and the signal/noise ratio of the estimated shock series declining as a result, identifying the impact of the shock has become harder.

To assess stability more systematically, we identify five subsamples based on the policy regime in place at the time following Bagliano and Favero (1998):

- 1969:1-1972:12—free reserves targeting
- 1973:1-1979:10—federal funds rate targeting
- 1979:11-1982:10—nonborrowed reserves targeting
- 1982:11-1988:10—federal funds rate-borrowed reserves targeting, pre- Greenspan
- 1988:11-2008:6—federal funds rate-borrowed reserves targeting, Greenspan /Bernanke period.[19]

Our first step is to analyze the stability of the regression coefficients via a series of Chow tests comparing each set of adjoining subsamples (Table 1). There appears to be some stability within the two post-82 subsamples (broadly corresponding to the great moderation period), but clear evidence of a structural break for the other potential break points. This suggests that Romer and Romer's reaction function, that assumes constant coefficients across the whole sample, could be misspecified.[20] These results are in line with those of Boivin and Giannoni (2006), who undertake a similar exercise for a small structural VAR similar to the systems discussed in Section II, and find strong evidence for a structural break between 1977 and 1986. Hence, the VAR identification methods discussed above—which like Romer and Romer's method assume time-invariant coefficients in the policy reaction function in order to identify monetary policy shocks—are likely to suffer from very similar problems.

**Table 1. Chow Stability Tests for Romer and Romer Policy Equation**

| | F test statistic | p-value |
|---|---|---|
| 69-72 vs. 73-79 | 2.44 | 0.003 |
| 73-79 vs. 79-82 | 7.65 | 0 |
| 79-82 vs. 82-88 | 4.89 | 0 |
| 82-88 vs. 88-08 | 1.77 | 0.029 |

Chow stability tests for structural breaks, using policy regime sub-periods identified in Bagliano and Favero (1998). F-test statistics robust to heteroskedasticity.

.Our second step is to test whether specific elements of $\Omega_t$ have changed. We focus in particular on two sets of variables: the eight forward-looking variables (1- and 2-quarter ahead forecasts) and nine backwards-looking variables (current and last quarter estimates) included in Romer and Romer's specification, and compare the post-1988 period with the rest of the sample. Table 2 presents F tests of the joint significance of the variables for the two subsamples. Policymaking appears to be unambiguously forward-looking in the post-1988 period, but one cannot reject the null hypothesis of no forward-looking variables in $\Omega_t$ during the pre-1988 period. This finding corroborates other analyses of Fed policymaking over the period.[21]

---

[19]We extend the last period from 1996:3 and start the first period in 1969:1 rather than 1966:1, reflecting the coverage of the original Romer and Romer series.

[20]Romer and Romer (2004) acknowledge the potential structural break around the 1979-82 period (actually, October 1979-May 1981), and show that their results are robust to dropping this particular subsample. However, we find that coefficients also differ significantly (with a p-value of 0.000) between the post-1982 sample and the pre-1979 sample, dropping the intervening period.

[21]For instance, Orphanides (2003) compares simple Taylor rules employing contemporaneous output gaps and inflation with forward-looking rules. While both types of rule appear to fit the data better in the post-1988 period compared with earlier periods, the contrast is more pronounced for the rule employing forecasts. Similarly, Boivin

(continued…)

**Table 2. Tests of forward and backwards-looking variables in Romer and Romer policy equation**

| | Forward-looking | | Backwards-looking | |
|---|---|---|---|---|
| | F Test statistic | p-value | F Test statistic | p-value |
| 1969:1-1988:10 | 1.48 | 0.169 | 3.22 | 0.001 |
| 1988:11-2008:6 | 3.90 | 0.000 | 8.00 | 0.000 |

F-tests of joint significance of 8 forward looking variables (quarters $q+1$ and $q+2$ ) and 9 backward-looking variables (quarters $q-1$ and $q$ ) in Romer and Romer's policy reaction function (see specification in Appendix Table A4). F-test statistics robust to heteroskedasticity.

These results shed some light on the findings presented in section II. Failure to allow for structural breaks—under all four methods of identification—will tend to give biased estimates of the shocks themselves, and hence biased estimates of the impact of the shock on other macroeconomic variables. For instance, by increasing the measurement error associated with the Romer and Romer shock series, it will lead to attenuation (bias toward zero) in the shocks' estimated macroeconomic impact.

The fact that policymaking appears to have become more forward looking in recent years has particularly serious implications for the VAR identification methods, since these do not include any forward-looking elements in $\Omega_t$ . As discussed in section I, if Fed policymakers react to an expected increase in output growth above the economy's potential by tightening monetary policy to partially offset it, then a monetary contraction will appear to cause higher growth if these anticipatory movements are not explicitly allowed for. Since anticipatory movements appear to have become more important for the recent period than earlier, this might explain why VAR identification methods identify the expected contractionary impact of monetary tightening for the earlier period, but for the later period generate the counterintuitive expansionary effects shown in section II. Although Romer and Romer's methodology attempts to control for anticipatory movements, by imposing equal coefficients throughout the sample it may not adequately capture the stronger effects in the recent period.[22]

These are unlikely to be the only misspecifications. For instance, all the identification methods above rely on the assumption that a relatively small number of variables adequately capture the Fed's information set. Since this is unlikely to be the case, omitted variable problems are likely significant (and may have become more pronounced in recent years as the Fed has made more

---

and Giannoni (2006) estimate a structural DSGE model that can account for the reduced responsiveness of the economy to monetary policy shocks since the 1980s uncovered by VAR analysis, and argue that the key explanation is a stronger Fed response to inflation expectations.

[22]However, if the changes to the parameters in the Fed's reaction function are due to changes in the Fed's preferences rather than in the transmission mechanism, then it is valid to ignore these when isolating policy shocks (because preference changes should be considered exogenous policy shocks and hence need to be included in the residual). The authors are grateful to David Romer for clarifying this point.

intensive use of a range of near-time indicators in its policy decisions).[23] In addition, the magnitude of monetary policy shocks has almost certainly been diminished by transparency-enhancing reforms to Fed communication practices since the early 1990s (Crowe and Meade, 2007), making it harder to identify the impact of shocks on the economy.

## IV. A NEW SHOCK MEASURE DERIVED FROM FED FUNDS FUTURES PRICES

### A. Overview

Conventional methods of identifying monetary policy shocks—which require the estimation of (1) with suitable proxies for $\Omega_t$—will perform badly if either $\Omega_t$ or $f()$ are misspecified. An alternative approach is to use financial market data to obtain the private sector's contemporaneous beliefs about $f(\Omega_t)$ at the time of each meeting, and use these as a proxy for the true reaction function and its elements. This circumvents the need to estimate $f(\Omega_t)$ directly, and therefore does not require that we impose restrictions on the variables in $\Omega_t$ or the functional form $f()$.

To illustrate this approach in general terms, assume that we have two measures of the private sector's expectation for the policy stance $S_t$ for a particular policy meeting: one in the immediate run-up to the meeting, ${}_{t-1}\widehat{S}_t$, and one immediately after the announcement of the policy stance decided at the meeting, ${}_{t}\widehat{S}_t$. Each is a noisy measure of the private sector's true expectation:

$$\begin{aligned} {}_{t-1}\widehat{S}_t &= E^P_{t-1}\left[S_t\right]+\xi_{t-1} = E^P_{t-1}\left[f\left(\Omega_t\right)\right]+\xi_{t-1} \\ {}_{t}\widehat{S}_t &= E^P_{t}\left[S_t\right]+\xi_{t} = S_t+\xi_t \end{aligned} \tag{7}$$

where the private sector's actual expectations at time $\tau$ of the stance at time $t$ are denoted by $E^P_\tau\left[S_t\right]$. The noise $\xi$ can arise from several sources, including time-varying risk premia as well as measurement or rounding errors. We make the following two identifying assumptions:

$$\begin{aligned} E^P_{t-1}\left[f\left(\Omega_t\right)\right]-f\left(\Omega_t\right)=0 \\ \xi_t-\xi_{t-1}=0 \end{aligned} \tag{8}$$

The first assumption states that the private sector's beliefs prior to the announcement about the

---

[23]An alternative methodology for incorporating the Fed's rich information set is the factor-augmented VAR approach (see, for instance, Bernanke and others 2005, and Bernanke and Boivin, 2003). One downside to this approach is that, even when one considers a wide range of potential variables, the Fed's information set—and the weights placed on different elements of it in the Fed's reaction function—are likely to change over time. It seems plausible that financial market participants have some useful information on these changes. Moreover, using this information rather than attempting to reconstruct the Fed's information set oneself is less data-intensive and allows for more parsimonious models. Identifying monetary policy shocks using Fed Funds futures market prices therefore offers a useful complimentary approach.

Fed's information set are correct.[24] The second assumption states that the noise term is unchanged around the time of the policy announcement. Then:

$$ {}_{t}\widehat{S}_{t} - {}_{t-1}\widehat{S}_{t} = s_{t} \tag{9} $$

This implies that a suitable proxy for the shock, $s_t$, is given by the change in the measure of the private sector's beliefs about the policy stance around the time of a policy announcement, ${}_{t}\widehat{S}_{t} - {}_{t-1}\widehat{S}_{t}$

## B. Fed Funds Futures Data

Our measures of the private sector's beliefs about the policy stance $\widehat{S}_t$ are derived from Fed Funds futures contracts. Our approach is similar to that in Kuttner (2001), Gürkaynak (2005) and Gürkaynak, Sack and Swanson (2005), although the details differ somewhat and these authors only look at the short term effect on financial variables rather than on the macroeconomy more generally.

The Federal Funds futures market was established at the Chicago Board of Trade (CBOT) in October 1988 (see Soderstrom, 2001; Kuttner, 2001 and Faust and others, 2004 for further information). The price of a contract for month $m+h$ (i.e. at a horizon $h$ from the current month $m$ ) is a bet on the monthly average effective Fed Funds rate in month $m+h$ (denoted $\bar{r}^{e}_{m+h}$ ). Note that the average target Fed Funds rate ( $\bar{r}_{m+h}$ ) might differ from the effective rate due to targeting errors on the part of the Fed:

$$ \bar{r}^{e}_{m+h} = \bar{r}_{m+h} + \varepsilon_{m+h} \tag{10} $$

These errors are typically small and mean zero. For a given contract price $p^h_d$ on day $d$ in month $m$, the futures rate $f^h_d$ is simply given by $1 - p^h_d$. Then standard no-arbitrage conditions imply that the futures rate is equal to the average effective Fed Funds rate in month $m+h$ , $E_d \bar{r}^{e}_{m+h}$ , plus a risk (or hedging or term) premium $\rho^h_d$:

[24]These assumptions are stated in their strongest form to clarify the exposition. A weaker assumption would be that, conditional on the realization of $\Omega_t$ and $s_t$, (8) holds in expectations. In this case, (E[s]) would also hold in terms of conditional expectations, but our proxy for $s_t$ could now include measurement error, leading to some attenuation bias when we use it for estimating the impact of policy shocks. Our strong identifying assumptions can be thought of as the limiting case, where in reality there could be some white noise terms on the right hand side of (8). As long as the variance of these error terms is relatively small, as seems likely given the short (24-hour) window around the policy announcement that we employ and the liquidity and competitiveness of the Fed Funds futures market, then the degree of measurement error should be limited. A more serious problem—simultaneity bias—will arise if (8) does not hold even in this weaker, conditional expectations, form, e.g. because the private sector makes systematic errors in forecasting the Fed's policy reaction function. This issue is addressed in more detail later in the paper.

$$f_d^h = E_d \overline{r}^e_{m+h} + \rho_d^h \tag{11}$$

Assuming that the risk premium $\rho_d^h$ remains constant and that there is also no change in the expected average targeting error $E_d\left[\varepsilon_{m+h}\right]$, then the change in the expected target rate during subsequent calendar months ( $h \geq 1$ ) following a policy announcement on day $d$ of month $m$ is given by:

$$\Delta E_d \overline{r}_{m+h} = f_d^h - f_{d-1}^h \tag{12}$$

while the change for the remainder of the current month (whose length is $M$ days) is given by:

$$\Delta E_d \overline{r}_m = \frac{M}{M-d}\left(f_d^0 - f_{d-1}^0\right) \tag{13}$$

The innovation to the expected target rate in a given month then serves as a good proxy for the underlying monetary policy shock $s_t$ under four assumptions. First, the average target rate $\overline{r}_{m+h}$ should be correlated with the policy stance $S_t$. If this holds then $f_d^h - f_{d-1}^h$ provides an estimate of ${}_t\widehat{S}_t - {}_{t-1}\widehat{S}_t$, while the noise term $\xi_t$ is given by the sum of the risk premium $\rho_d^h$, the expected Fed targeting error $E_d\left[\varepsilon_{m+h}\right]$ as well as data errors. Second, there should be no predictable changes in the noise terms that make up $\xi$, e.g. due to predictable effects of policy announcements on risk premia: this is a necessary condition for the second assumption in (8) to hold. Third, there should be no other 'news' that might affect the expected futures rate (such as macroeconomic data announcements that might have implications for rate changes in the future) during the 24-hour period associated with the policy decision. Last, the policy announcement itself should not reveal information about the Fed's private information set $\Omega_t$ or its reaction function $f()$ . These last two assumptions are necessary for the first assumption in (8) to hold.[25] Assuming that these assumptions are valid, then the policy 'surprise' is a good measure of the shock. The evidence, discussed in section V, provides strong support for the first three assumptions, while evidence on the fourth is more mixed.

---

[25]For instance, a negative macroeconomic news release that occured concurrently with a policy announcement would imply lower rates in the future, *ceteris paribus*. Hence, conditional on this new information (an element of $\Omega_t$ ), expectations relating to the systematic component of the policy stance before the meeting would have been too high, and (8) is contradicted. Similarly, if a policy announcement provides new information about the Fed's information set, e.g. so that a rate cut signals that the Fed expects a recession, then the private sector's beliefs prior to the announcement were incorrect and again (8) does not hold.

## C. Constructing the Shock Series

Our analysis focuses only on FOMC meeting dates, rather than on all dates that the Fed announced changes to the target Fed Funds rate, including inter-meeting changes. We choose this strategy for several reasons. We believe that decisions to *not* change rates—when a rate change might have been expected by the private sector—also constitute monetary shocks. If one did not limit attention to FOMC meeting dates, then for consistency one would have to consider every day as one when rates could have been changed. But in this case it becomes difficult to identify monetary policy shocks, because for most days other sources of news are more likely to account for any change in futures rates than the lack of a rate announcement.[26]

The simplest signal of the policy stance $S_t$ is the futures rate for the current month, $f_d^0$.[27] However, we argue that there are several reasons why innovations to futures rates further along the maturity structure offer additional information about the shock which can be usefully incorporated. First, all the innovations will include some noise, including due to changes in the risk premium, changes in beliefs about targeting errors (i.e. persistent deviations of the effective Fed Funds rate from the target) and rounding errors. Hence, combining the information from several sources—essentially taking a sample mean of the shock measures obtained from contracts at different horizons—should help to minimize the effect of these errors to the extent that they are idiosyncratic across the innovations at different horizons. This averaging may be particularly important since the risk premium is likely to be more volatile at shorter horizons (as we show in the data appendix, the market for the current month contract is not the most liquid, and intra-month trading volumes are in fact particularly volatile for this contract, which could lead to a more volatile liquidity premium and hence introduce more noise into the shock measure). In fact, data on trading volumes indicates that no single contract is traded on every day that a policy announcement is made, whereas there is always trading in contracts at two or more maturities on such days. Assuming that prices on actively trading securities are likely to provide a better gauge of expectations, this points to a clear benefit in combining information from contracts at various maturities, rather than relying on a particular maturity.

Moreover, since the Fed's policy decisions are relatively persistent over time, a policy change in

---

[26]If we were to include only non-meeting days when rates actually change, we might incorporate some additional information on shocks, but at the expense of biasing our sample, since the decision to change rates outside of the regular meeting schedule is likely to be non-random. Because rate changes on FOMC meeting dates are relatively common, while rate changes outside FOMC meeting dates are relatively rare (particularly after 1991), it seems to us that in focusing on the FOMC meeting dates only we do not lose a significant quantity of information. Moreover, like Faust and others (2004), who come to the same judgment, we believe that intermeeting changes are more likely to be associated with the simultaneous release of macroeconomic information rather than reflecting exogenous shocks to policy. Hence, these observations are likely to provide only noisy information on the monetary policy shock associated with the rate decision.

[27]This is the approach followed by Kuttner (2001). Since the scaling factor can become very large in the last few days of the month, amplifying any noise in the shock measure, Kuttner uses the innovation to the next month's futures contract as a proxy in these cases. This is still a good measure of the surprise element of the rate change at the current meeting since the meeting schedule implies that an FOMC meeting late in the month means no meeting during the subsequent calendar month. We follow the same methodology for our current month shock measure.

the current period will be reflected in higher expected rates several months ahead, so that futures contracts settling several months in the future will also contain information about the current shock. Indeed, shocks which are expected to be permanent might be expected to have a greater impact on the economy. But some shocks to current rates might have little impact on longer term expectations (for instance, if the shock were to the immediate timing of the rate change rather than to the long-term direction of rates, as Gürkaynak, 2005, argues). Hence, a measure of shocks that combines the innovations to rates in the current (spot) month with those anticipated in the future is likely a better measure of the overall policy stance.[28]

While contracts are now available for more than a year into the future, longer-dated contracts have not been available for the whole period and even now are typically relatively illiquid. Hence, we focus on contracts for the current month and up to 5 months ahead. In order to combine the information available in the estimated forecast innovations at all six horizons, we estimate a factor model via maximum likelihood. We find that two factors adequately capture the information in the futures shocks.[29] Table 3 displays the factor loadings and unique variances. Table 4 displays the correlation matrix for the two factors, the individual shocks for the six monthly contracts, and the change in the actual Fed Funds target rate.

The two factors summarize the new information on the medium term evolution of policy rates that is revealed by the policy rate announcement. Indeed, the factors turn out to have an intuitive interpretation. The first factor, which is highly positively correlated with all the individual innovations, can be thought of as a levels effect: that portion of the new information related to the policy announcement that causes vertical shifts in the expected medium-term trajectory for policy rates. Since the transmission of monetary policy is generally thought to occur via the impact of short rate changes on longer term (real) rates, it is this portion of the new information on rates that corresponds most closely to the relevant policy shock. We therefore use this factor as our measure of the underlying policy shock.

[28]In fact, when we attempt to replicate our baseline results using the Kuttner-style spot month futures rate innovation, rather than our prefered measure that combines information across several futures contracts, we obtain perverse IRFs similar to those obtained using VAR and narrative identification schemes. This appears to validate our approach. There is further discussion of these results in section V.

[29]Estimating a principal factor model with up to six factors, the first factor accounts for 92 percent of the total variance, the second factor for a further 9 percent, and the third factor for 0.4 percent. The eigenvalues of the first three factors are 5.2, 0.52 and 0.02 respectively (the last three factors have negative eigenvalues and make a cumulative contribution to the variance of -1 percent). Hence, a model with two factors appears to adequately and parsimoniously capture the main patterns of correlation in the data, and it is this parsimonious specification that is then estimated via Maximum Likelihood.

## Table 3. Factor analysis: factor loadings and unique variances

| Horizon (+ months) | Factor 1 | Factor 2 | Uniqueness |
|---|---|---|---|
| 0 | 0.744 | 0.562 | 0.131 |
| 1 | 0.884 | 442 | 0.024 |
| 2 | 0.972 | 0.121 | 0.04 |
| 3 | 0.987 | -0.034 | 0.025 |
| 4 | 0.985 | -0.142 | 0.01 |
| 5 | 0.958 | -0.223 | 0.032 |

Factor loadings and unique variances obtained via Maximum Likelihood factor model with two factors imposed, estimated over 157 per-meeting observations (December 1988-June 2008).

## Table 4. Correlation Matrix, Fed Funds Futures shocks

| | Rate Change | Current | +1 mth | +2 mths | +3 mths | +4 mths | +5 mths | Factor 1 | Factor 2 |
|---|---|---|---|---|---|---|---|---|---|
| Rate Change | 1 | | | | | | | | |
| Current | 0.34 | 1 | | | | | | | |
| +1 mth | 0.35 | 0.91 | 1 | | | | | | |
| +2 mths | 0.38 | 0.78 | 0.91 | 1 | | | | | |
| +3 mths | 0.39 | 0.71 | 0.85 | 0.97 | 1 | | | | |
| +4 mths | 0.38 | 0.66 | 0.81 | 0.94 | 0.98 | 1 | | | |
| +5 mths | 0.34 | 0.59 | 0.75 | 0.9 | 0.95 | 0.98 | 1 | | |
| Factor 1 | 0.39 | 0.75 | 0.89 | 0.97 | 0.99 | 0.99 | 0.96 | 1 | |
| Factor 2 | 0.04 | 0.58 | 0.46 | 0.13 | -0.04 | -0.15 | -0.23 | 0 | 1 |

Correlation coefficients: variables are the change in Fed Funds target rate, the monthly shock measures outlined in the main text (current month through 5 months ahead) and the first and sector factors obtained via Maximum Likelihood. Factor 1 is our shock measure. Estimated over 157 per-meeting observations (December 1988-June 2008).

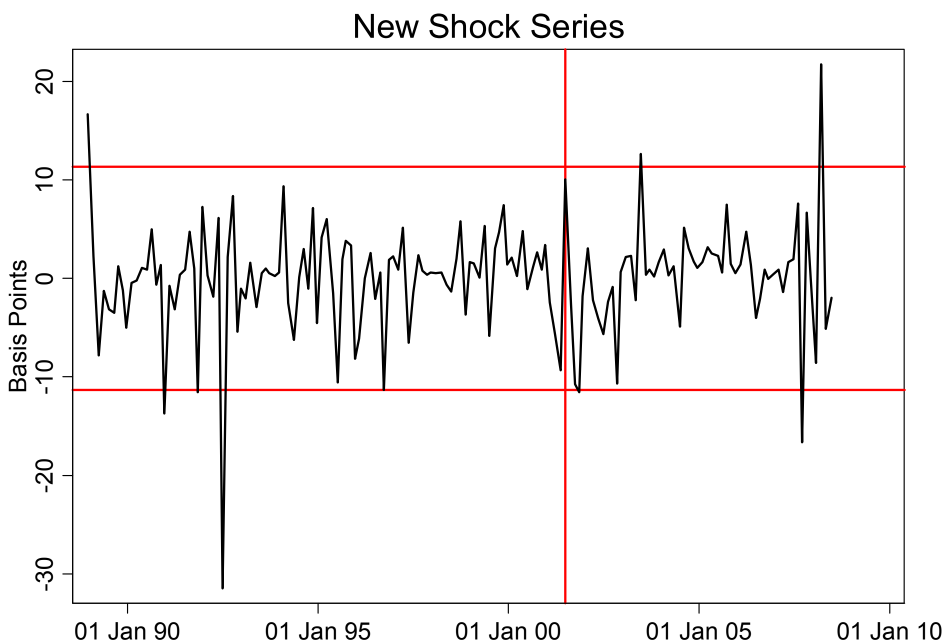
The second factor, whose correlation with the individual innovation series at different maturities decreases monotonically from positive to negative as the maturity increases, can be thought of as a slope or yield curve effect: that portion of the new information relating to the policy announcement that leads to differential effects on expected policy rates in the near term and further out. While this factor captures an important portion of the news relating to policy announcements, it does not capture the notion of a policy shock that is the focus of the current paper.[30]

[30]Hamilton (2008) argues that the slope factor (derived from a three factor model of the first three Fed Funds futures contracts) is key in explaining movements in mortgage interest rates. However, in keeping with our prior, we find little evidence that our estimated slope factor has any significant impact on output or prices (results available from the authors).

### D. Assessing the Shock Series

Our new shock series is presented in Figure 6. Our factor-based shock measure has a mean of 0 and a standard deviation of 1 by construction. To aid interpretation, in Figure 6 it is scaled to be a weighted average of the deviations from the mean of the six underlying monthly shock series. Two standard deviation bars are shown, and the June 27 2001 meeting is indicated by a vertical bar to aid the discussion in sub-section E.

**Figure 6. Time Series of New Shock Measure**




New shock series, in basis points. To make it comparable in size to the 6 underlying shocks, the first factor (SD=1 by construction) is divided by the sum of the 6 coefficients from the factor model. Two standard error bands shown by horizontal lines; vertical line identifies the June 2001 FOMC meeting discussed in Section IV part E.

The validity of our shock measure depends on the validity of the underlying assumptions. Our first assumption, that the Fed Funds target rate at the relevant horizons (0-5 months) is correlated with the 'true' monetary stance, seems uncontroversial. Bernanke and Mihov (1998) have demonstrated that a Fed Funds targeting model best describes monetary policy in the post-1988 period, while it is intuitive that, in an economy with forward-looking agents making irreversible economic decisions, the overall stance of policy depends not only on the current target rate but also on the rates expected in the immediate future. With respect to our second assumption—that there should be no predictable innovations to the noise component of the private sector's expectations about the policy stance in the short run—Piazzesi and Swanson (2008) show that anticipated changes to risk premia in the Fed Funds futures market occur mainly at business cycle frequency. With respect to our third assumption—that other information that could be conflated with the policy announcement and bias our results is not released on the same day—Gürkaynak and others (2005) show that some FOMC meeting and intermeeting dates associated with policy announcements coincide with macroeconomic data releases. However, they show that only *Employment Report* releases have any independent effect on Fed Funds futures. Bernanke and Kuttner (2005) identify ten observations, all before 1994, for which *Employment*

*Report* releases coincide with policy announcements or FOMC meetings. But our decision to focus only on FOMC meetings helps to alleviate this problem, since only three of these dates coincide with FOMC meetings (the others coincide with intermeeting changes).[31] We provide some empirical evidence that the inclusion of these dates is not driving our results in the robustness checks in section V.

To test our fourth assumption, we regress our (scaled) shock measure on the difference between the Fed's Greenbook forecasts and high-quality private sector (*Blue Chip*) forecasts for the 17 variables used in Romer and Romer's (2004) estimated reaction function, where this difference is used as a proxy for the Fed's private information. Since the Greenbook forecasts are only made public with a 5-year lag, the shock measure should only be correlated with the Fed's private information to the extent that the latter is revealed indirectly by the policy rate, the announcement and any related communication. As we show in Table 5, the joint hypothesis of zero coefficients on all 17 variables cannot be rejected at the 10 percent level. This suggests that our shock measure should be relatively uncorrelated with the Fed's private information, and simultaneity bias should therefore not be a significant problem.

However, an inspection of the coefficient estimates in Table 5 points to evidence that our shock measure may be contaminated by the impact of the Fed tightening policy in response to near term output and price pressures, since our shock measure responds positively to current quarter output and inflation forecasts. We investigate further the implications of this for our results in section V.

To illustrate how our shock measure compares to others in the literature, Table 6 presents correlation coefficients for our shock measure (New), the change in the target Federal Funds rate $\Delta FF$ and Romer and Romer's shock measure (*R&R*; all on a per-meeting basis, for 157 meetings); the final row presents correlation coefficients between the per-quarter average of these three measures and the monetary policy shock obtained from a Cholesky decomposition of Christiano and others' quarterly VAR specification (*CEE*), for 76 quarterly observations (1988Q4-2007Q3). Our new shock measure is positively and significantly correlated with all three measures (at least at the 10 percent level).

[31] The three dates in question are 7 July 1989 and 2 July 1992 (the day after the meeting), and 4 February 1994 (the day of the meeting).

**Table 5. Regression results and F-test statistics for policy shock measure and Greenbook**

| | |
|---|---|
| $\text{Unemployment}_0$ | -4.26 |
| $\text{Output Growth}_{-1}$ | -1.31 |
| $\text{Output Growth}_0$ | 2.37*** |
| $\text{Output Growth}_1$ | -0.783 |
| $\text{Output Growth}_2$ | 1.19 |
| $\text{GDP Deflator}_{-1}$ | -0.92 |
| $\text{GDP Deflator}_0$ | 2.34** |
| $\text{GDP Deflator}_1$ | -1.49 |
| $\text{GDP Deflator}_2$ | -0.323 |
| $\Delta\text{Output Growth}_{-1}$ | 0.541 |
| $\Delta\text{Output Growth}_0$ | -1.14 |
| $\Delta\text{Output Growth}_1$ | 0.803 |
| $\Delta\text{Output Growth}_2$ | -1.44 |
| $\Delta\text{GDP Deflator}_{-1}$ | 0.300 |
| $\Delta\text{GDP Deflator}_0$ | -1.31 |
| $\Delta\text{GDP Deflator}_1$ | -0.117 |
| $\Delta\text{GDP Deflator}_2$ | 1.22 |
| Constant | -0.610 |
| $R^2$ | 0.185 |
| F(17) | 1.50 |
| p-value | 0.132 |

The dependent variable is the scaled shock measure in basis points; the independent variables are the difference between the Greenbook and Blue Chip forecasts for the 17 variables identified by Romer and Romer (variables are estimates for the previous or current quarter or forecasts one or two quarters ahead, except for variables denoted Δ which are the change in the forecast from the previous meeting; all variables are then differenced between the Greenbook and Blue Chip consensus forecasts). The regression is run over 113 FOMC meetings between 1988 and 2002. The F-test statistic shown is for the joint null hypothesis that the coefficient on all 17 variables is zero. Standard errors are robust to heteroskedasticity (but are omitted from the table for brevity). Significance levels indicated by *** (1 percent); ** (5 percent); * (10 percent).

## E. Our New Shock Series: An Illustrative Observation

Our shock measure, although correlated with existing measures, can differ significantly from these for some observations. These differences can help illustrate some of the relative strengths (and weaknesses) of our approach. For instance, the FOMC decided at its June 26-27 2001 meeting on a 25 basis points reduction in the Fed Funds rate. The cut followed five successive 50 basis point cuts (three at the three preceding meetings and two cuts between meetings), as part of a rate-cutting cycle that saw the Fed Funds rate fall from 6.5 to 1.75 percent over the course of the year.

**Table 6. Correlation between Shock Measures**

| | New | ΔFF | R&R | CEE |
|---|---|---|---|---|
| New | 1 | | | |
| ΔFF | .39*** | 1 | | |
| R&R | .23*** | .73*** | 1 | |
| CEE | .22* | .26** | 0.09 | 1 |

Correlation coefficients for our new shock measure (*New*) and existing measures: the change in Fed Funds Rate $\Delta FF$ , Romer and Romer's narrative measure (*R&R*), and Christiano and others' measure (*CEE*; based on Cholesky decomposition of VAR residuals). Coefficients in rows 1-3 based on 157 per-meeting values; coefficients in last row based on 76 quarterly values. Significance levels indicated by *** (1 percent); ** (5 percent); * (10 percent).

How do different shock measures treat this observation? The monetary policy shock from Christiano and others' system for the second quarter of 2001 is significantly negative (although this also reflects the 50 basis point rate cut at the May meeting and a similar inter-meeting cut in April). However, it is clear from reading the Fed's statements as well as from market reaction that the Fed's interest rate cuts were largely in response to the economic slowdown in the wake of the bursting tech bubble and concerns that the economy was set to slow further. For instance, the statement accompanying the decision states that "the risks are weighted mainly toward conditions that may generate economic weakness in the foreseeable future." This example illustrates the failure of conventional VAR identification schemes to convincingly address the endogeneity problem. Romer and Romer's method conditions on the Fed's private information, but this shock measure is also negative—perhaps because the reaction function underestimates the Fed's response to forward-looking information as argued in section III.

By comparison, our shock measure is large and positive (almost 2 standard deviation bands above zero, or 10 basis points when suitably scaled). The intuition for this is that market participants were anticipating another 50 basis point cut in rates. Reuters (June 28) reports that "the market had priced in the prospect for 50 basis points." The smaller cut therefore represented a positive shock to Fed Funds rate expectations. Market reaction to the move supports our interpretation of the June 27 rate cut as a policy tightening. Reuters (June 27) reports that "the dollar climbed to a 10-week high on the yen on Thursday, helped by a raft of factors, including the... rate cut." The dollar also gained ground against the euro. Meanwhile, bond yields rose significantly (particularly for two-year government paper). These reactions are more consistent with a contractionary than an expansionary monetary policy shock.

However, market reaction also illustrates the difficulty in disentangling the random "shock" component of the policy news from the potential revelation of the Fed's private information. For instance, Reuters (June 28) reports analysts' belief that "the slower pace of easing could signal the Fed thinks the economy's state is not dismal." Similarly, it argues that "stocks are expected to have a firm start... as investors cross their fingers that the Federal Reserve's smaller-than-hoped-for drop in interest rates is a hint that the nation's economy is not in dire straits." These quotations suggest that the private sector updates its expectations about the economy based in part on what it can infer from the policy decision about the Fed's private information (as Romer and Romer, 2000 confirm). This creates an identification problem: if our shock measure captures the Fed's private information, the estimated impact on economic outcomes will be contaminated by simultaneity bias. However, because market participants cannot assign the Fed's decision to its private information with any certainty, we would argue that the policy surprise on June 27 2001 can still be characterized primarily as a policy shock. Section V discusses this issue in more detail.

## V. Identifying the Effect of Monetary Policy Shocks Using Our New Measure

### A. Baseline impulse responses and forecast error variance decompositions

Following Romer and Romer, we identify the effect of monetary policy shocks using a small 3 variable VAR. The variables are ordered so that monetary policy is allowed to respond to, but not affect, output and inflation contemporaneously. We use the log of industrial production as our output measure and the log of the consumer price index as our measure of prices.[32] As with Romer and Romer's shock measure, our measure captures unanticipated *changes* in policy rates. Hence, like Romer and Romer, we enter our shock measure cumulated in the VAR, since here it is the level, not the change, in policy that is the appropriate variable.[33] To facilitate comparison with Romer and Romer's VAR results, the baseline VAR includes 36 monthly lags. However, the results are fully robust to shorter lag specifications that match the kind of lag structure in the other VAR results cited in Section II and make less demands on the data given the relatively short sample available. Results for 6, 12 and 24 months, which are almost identical to the baseline impulse responses, are shown in the Appendix.

Impulse response functions are shown in Figure 7. After almost one year, a contractionary monetary policy shock shows a sustained negative effect on output that has its maximum impact at a horizon of around two years. The output response is very similar to the baseline results for the earlier period reviewed in section II (although with greater persistence), but very different from the results obtained for the 1988-2008 period using the same methodologies.

The response of prices to a monetary contraction is more problematic. The effect becomes significantly negative only after four years; the positive response over the medium term, although small, contrasts with the negative effect that has generally been found in the literature. Castelnuovo and Surico (2006), like Hanson (2004), find evidence that the price puzzle is limited to the pre-1979 period, arguing that this is due to equilibrium indeterminacy when monetary policy responds weakly to inflation expectations, and that the inclusion of a variable capturing the persistence of expected inflation under indeterminacy can eliminate the price puzzle. However, our baseline results suggest evidence for a price puzzle even in the post-Volcker

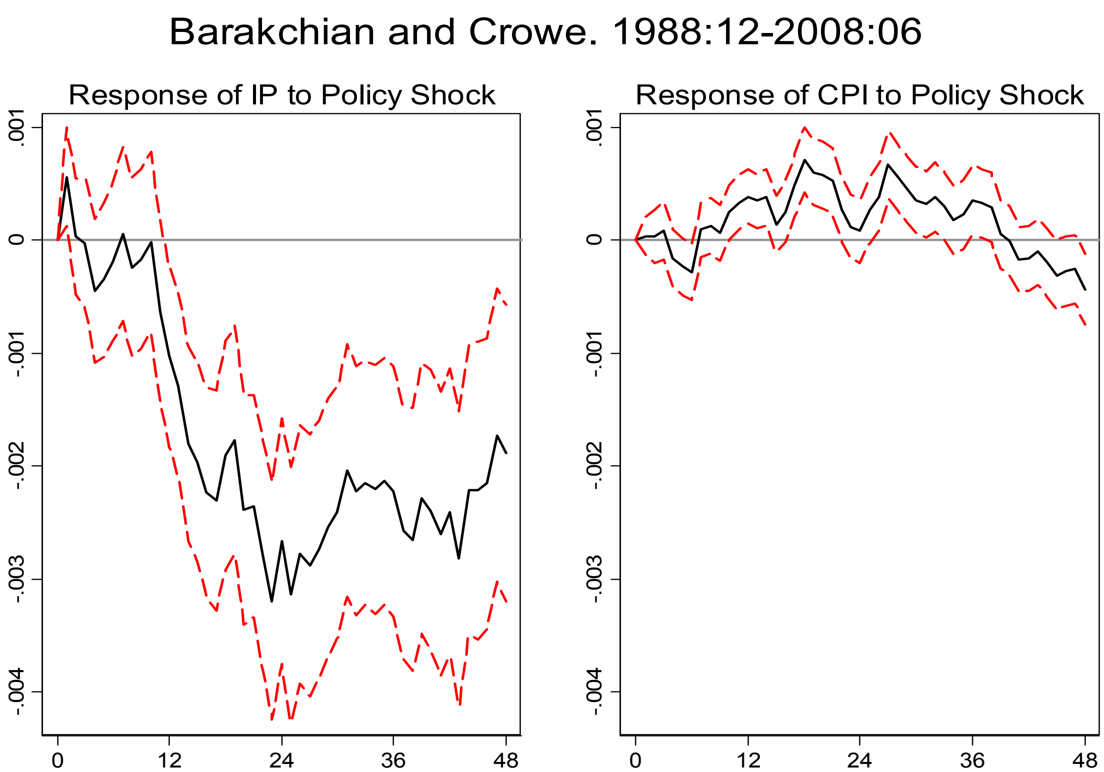
---

[32]This follows much of the literature, but differs from Romer and Romer (2004) who use the log of the producer price index for finished goods as their price measure. Our VARs also include an exogenous time trend.

[33]An additional rationale for using the cumulated shock series, which is I(1) by construction, is that the output and price series are generally considered I(1); hence, if the I(0) shock series were included the VAR would be statistically unbalanced, leading to nonstationary, highly persistent, residuals. Including the I(1) cumulated series allows for implicit cointegration between the variables in the VAR.

period, when the reaction of interest rates to expected inflation should be sufficiently strong to guarantee equilibrium determinacy. Other studies (e.g. Christiano and others 1996) have included a measure of commodity prices as a means of eliminating the price puzzle (although their argument for including this variable, that commodity prices help to forecast inflation, has been criticized by Hanson, 2004).[34] In the following section we add a proxy for inflation expectations and a commodity price index to our baseline VAR as two of a series of robustness checks; neither helps to resolve the price puzzle. However, this apparently robust finding of a significant price puzzle is consistent with other recent work that uses Fed Funds futures to identify policy shocks (Thapar, 2008).

**Figure 7. Impulse Response Functions**



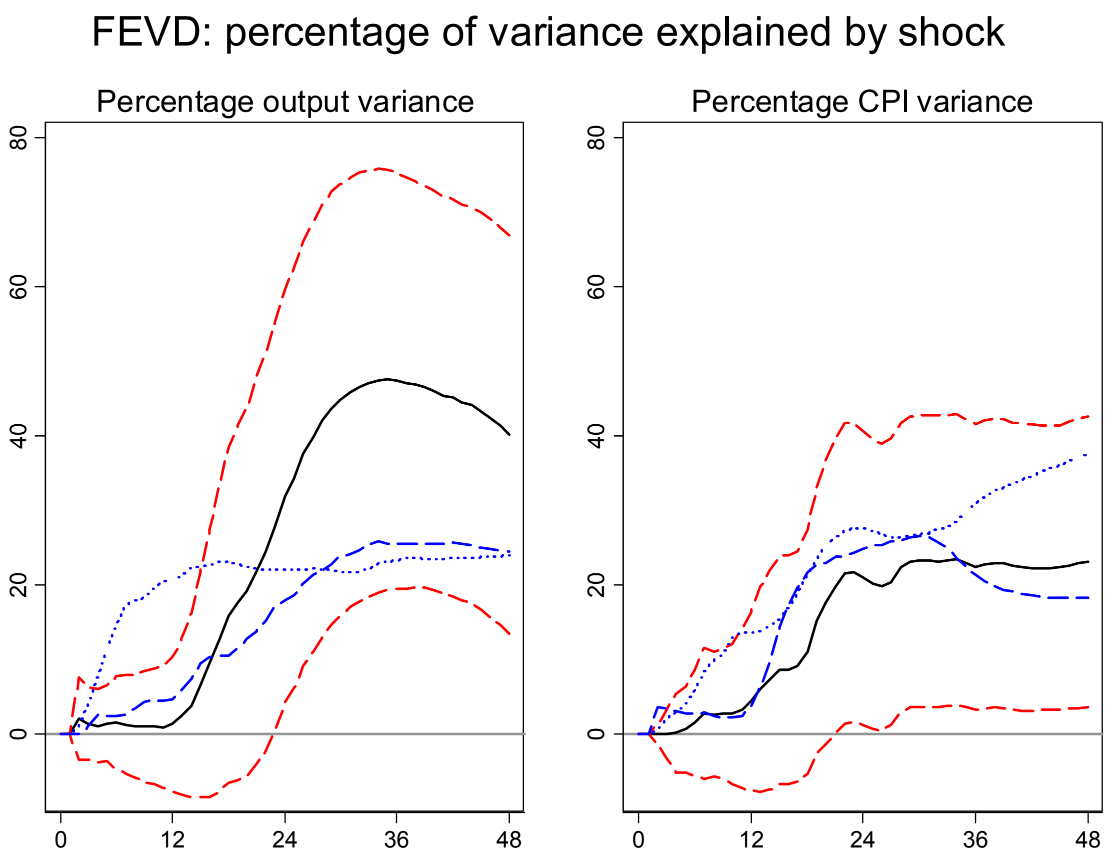

Structural VAR (Monthly data, 3 endogenous variables plus constant and linear time trend, 36 lags).
Variables ordered as industrial production, consumer prices, both seasonally adjusted and in logs, and our shock measure, cumulated. Graphs show response of industrial production and CPI to a one standard deviation positive shock to the policy measure.
Structural shocks obtained via Cholesky decomposition.
Two Standard Error bands produced by parametric bootstrapping (500 replications).

Received wisdom about the great moderation period is that less pronounced monetary policy shocks have helped to contribute to the general moderating in macroeconomic volatility. In order to shed some light on this issue, we analyze the percentage of the forecast error variances of

[34]Giordani (2004) argues that the price puzzle arises because the VAR system, by including output rather than the output gap (which enters in theoretical models), is misspecified. However, since our VAR model includes a linear time trend, we are in effect dealing with an output gap measure, assuming that (log) potential output follows a linear trend. This explanation is therefore unlikely to account for the estimated price puzzle in our model.

output and prices which can be attributed to our shock measure and two other measures over the recent period, a Federal Funds rate shock and the Romer and Romer shock (Figure 8).[35] Results for our shock measure are shown with a solid line; results for Fed Funds rate shock (dashed line) and Romer and Romer shock (dotted line) are shown for comparison; two standard error bands for our shock measure are also shown.

**Figure 8. Forecast Error Variance Decomposition for New Shock Measure**




Structural VAR (Monthly data, 3 endogenous variables plus constant and linear time trend, 36 lags). Variables ordered as industrial production, consumer prices, both seasonally adjusted and in logs, and one of three policy measures: our shock measure; Romer and Romer's measure (both cumulated); and the Federal Funds rate. Graphs show Cholesky FEVDs: the percentage of the forecast error for output and CPI accounted for by each policy measure. The FEVD for our shock measure is shown in bold, with two standard error bands produced by parametric bootstrapping. FEVDs for the Fed Funds rate (dashed line) and Romer and Romer shock (dotted line) are shown for comparison (SE bands not shown).

---

[35]In order to make the results comparable, we estimate in each case a small recursive VAR including industrial production, CPI and one of three variables: the Federal Funds rate, the Romer and Romer (cumulated) shock measure and our (cumulated) shock measure. The sample period is 1988:12-2008:06. This approach is similar to that of Romer and Romer (2004), who estimate the first two VARs to compare impulse responses using their shock measure with those using a standard recursive VAR shock measure (with the Fed Funds rate as the monetary instrument). However, we use CPI as our price measure, whereas Romer and Romer use the PPI for finished goods.

The estimated impact of monetary policy shocks on the variance of the price level is broadly similar across the three measures, although the Romer and Romer method identifies the largest effect, particularly at longer horizons, which is intuitive given the impulse response shown in Figure 4. However, at horizons of more than two years the estimated impact on output volatility is considerably higher for our shock measure—around 2 times as high as with either of the alternative measures. In fact, the results using our new measure suggest that almost half of forecast error variance at horizons of around 3 years can be accounted for by monetary policy shocks. Hence, while monetary policy shocks may have moderated in absolute terms, their relative contribution to output volatility in recent years may need to be reassessed.

### B. Robustness

We report here results for eight robustness checks and one further comparison. First, we change the ordering in our baseline VAR, allowing our monetary policy shock to have an instantaneous impact on output and prices. Impulse responses remain qualitatively identical, although the price puzzle is more pronounced. Second, we use Romer and Romer's price measure (PPI for finished goods). Again, the only (modest) difference is in the strength and persistence of the price puzzle. Third, we modify the lag structure to include 6, 12 or 24 lags rather than 36. The estimated impulse responses are essentially unchanged. Fourth, we assess subsample stability by estimating the baseline VAR (with lag length reduced to 12 in light of the shorter sample) for two truncated time periods. The first drops the period up to and including the 1990-91 recession, the second drops observations from the 2001 recession onwards, using recession dates from the NBER. Results are qualitatively identical to those for the sample as a whole.[36]

As a fifth robustness check, we include a commodity price index, ordered first in the recursive VAR. As already discussed, this has helped to eliminate the price puzzle in some studies. However, we find that including commodity prices has a similar effect to using the PPI price measure (i.e. the price puzzle remains); in addition, positive (contractionary) monetary policy shocks are associated with significant increases in commodity prices, while the output response to the policy shock is unchanged.

As a sixth exercise, we include a measure of inflationary expectations to test the robustness of the price puzzle. Following Castelnuovo and Surico (2006), we use one quarter ahead expected inflation from the Fed's Greenbook (replaced by the corresponding *Blue Chip* forecast for 2003 onwards), and order this variable first in the recursive VAR. This exercise does not help to eliminate the price puzzle either, and the output response is also unaffected.

As a seventh robustness check, we assess whether the inclusion of FOMC meeting dates that coincide with Employment Report releases is critical to the results. We run the baseline VAR with dummy variables that take a value of one for each of the months in question and zero

[36]As an additional robustness check we started the sample in August 1993 so as to drop the large negative shock in July 1992 (the model is estimated with 12 lags). Results are identical to the baseline, indicating that this observation is not driving the results (results of this exercise are available from the authors).

otherwise (entered as exogenous variables in the VAR). The VAR is estimated with 36 lags of the endogenous variables and each dummy enters contemporaneously and with 12 lags.[37] The output response to the policy shock remains the same as under the baseline; the modest price puzzle remains in the medium term, but the negative price effect after 3-4 years becomes more pronounced. Because our shock measure is identified outside the VAR it seems likely that our results are robust to other modifications to the VAR framework.

Finally, we estimate single-equation systems for output and prices similar to those estimated by Romer and Romer (2004). In keeping with the VAR results, we find a negative and persistent effect on output (with a point estimate of between 1 and 2 percent) and a small positive effect on the price level (although, due to wide estimated standard error bands, both effects are only at the border of statistical significance).[38]

This section closes with a final comparison exercise. To shed some light on how our factor-derived shock measure compares with the simple Kuttner (2001) spot-month shock, we estimate the baseline model with the (cumulated) spot-month innovation in place of our shock measure. In this case, the impulse response for output is closer to that for the other identification schemes, with a small, albeit insignificant, positive output response to a 'contractionary' policy shock. These results support the view that shocks to the spot month futures contract mainly reflect new information about the timing, rather than the general direction, of policy. Hence, it is not surprising that the IRFs associated with this noisy shock measure are imprecisely measured. As with the other identification schemes discussed in section 2, the apparently perverse sign of the estimated effect of policy on output is suggestive of simultaneity bias, perhaps because timing shocks are particularly associated with the Fed's communication of private information. This interpretation seems plausible given that the Fed's information advantage is generally thought of as being most pronounced for the kind of near-term indicators that might be of particular relevance for the decision to bring forward or postpone an anticipated rate change decision.

### C. Decomposing our Shock Measure

In section IV we presented evidence that our shock measure may be contaminated by the Fed's reaction to private information on near term inflationary pressure, with the Fed reacting to perceived higher output growth and inflation by tightening policy, as one might expect given its price stability mandate. In terms of the 2-equation model of the economy given in (2), these results imply that $\varphi > 0$, so that our baseline results may be biased (toward zero in this case), as shown in equation (4).

Romer and Romer's (2000) analysis of Fed and private sector forecasts suggests that the Fed's

---

[37]We identify three months with contemporaneous *Employment Report* releases and FOMC announcements, but the dummy for the first such month drops out once the sample is adjusted to accomodate the lag structure.

[38]Results of all robustness checks are shown at the end of the accompanying appendix.

forecasts are likely to include some accurate private information. To shed some additional light on this issue, we regress the Fed's private information (the difference between the Fed's forecast and the private sector forecast) on the private sector's overall forecast error (the difference between the actual outcome and the private sector forecast), for both real GDP and the GDP deflator and at forecast horizons of 0 to 2 quarters. The $R^2$ s from these regressions have the interpretation of the share of the Fed's private information that turns out to be correct ex post. This share varies from 1 to 6 percent for real GDP and from 3 to 20 percent for the GDP deflator (results reported in the Appendix, Table A2). In terms of our simple 2-equation model of the economy, the results suggest that $Cov\left(\widehat{u}_t, u_t\right) > 0$, again pointing to positive bias in our estimate of the effect of policy on output and inflation. Note though that the Fed's accurate information accounts for a relatively small share of the difference between its forecast and the private sector's, suggesting that the bias in (4) is small.

To provide some additional evidence on the likely impact of simultaneity bias on our results, we decompose our shock measure using the results of the regression of the shock on the Fed's private information presented in Table 5. The residuals $\widehat{S}_t^r$ from this equation give an estimate of the 'pure' shock component $s_t$, while the fitted values $\widehat{S}_t^f$ give an estimate of any remaining portion of the systematic component $f\left(\Omega_t\right)$. However, simultaneity bias is not the only likely source of bias in our results. Attenuation bias (bias towards zero) due to measurement error is also likely to be present, although we would expect its magnitude to be relatively small.[39] While the residual $\widehat{S}_t^r$ should be cleansed of simultaneity problems, if $f\left(\Omega_t\right)$ is correctly specified then the fitted value $\widehat{S}_t^f$ will be cleansed of measurement error (it will all be captured by $\widehat{S}_t^r$).[40] Comparing impulse responses from these two orthogonal portions of our overall shock measure therefore provides an indication of the relative magnitude of the biases created by simultaneity and measurement error.

We enter the two decomposed shock measures in our baseline VAR system (with $\widehat{S}_t^f$ and $\widehat{S}_t^r$ ordered third and fourth, respectively; Figure 9).[41] Both the 'predicted' portion of the shock (bottom panels) and the residual portion (top panels) have a significant negative effect—of strikingly similar magnitude—on output. This suggests that, for output, simultaneity bias is of around the same order of magnitude as the bias due to measurement error, where this latter bias is likely to be small. Moreover, since both sources of bias should bias the estimated effect towards zero the true effect is likely somewhat larger. Note that the fitted portion of the shock

[39]Measurement error could enter from the estimation of the factor model, from time variation in the risk, hedging and liquidity premia associated with the futures contracts as well as from rounding and data errors.

[40]This is the logic behind instrumental variable (IV) estimation as a solution to measurement error. See Greene (2008) pp. 325-7.

[41]Since we lose data after 2002 and there is an additional variable in the VAR, we adopt a more parsimonious lag structure (12 lags).

measure accounts for 17 percent of output variation at a 3 year horizon, while the residual portion accounts for 30 percent, reflecting the fact that most of the variation in our shock measure cannot be accounted for by the Fed's private information. Finally, we find that most of the price puzzle appears to stem from the residual portion of the shock, suggesting that simultaneity bias is not its source.

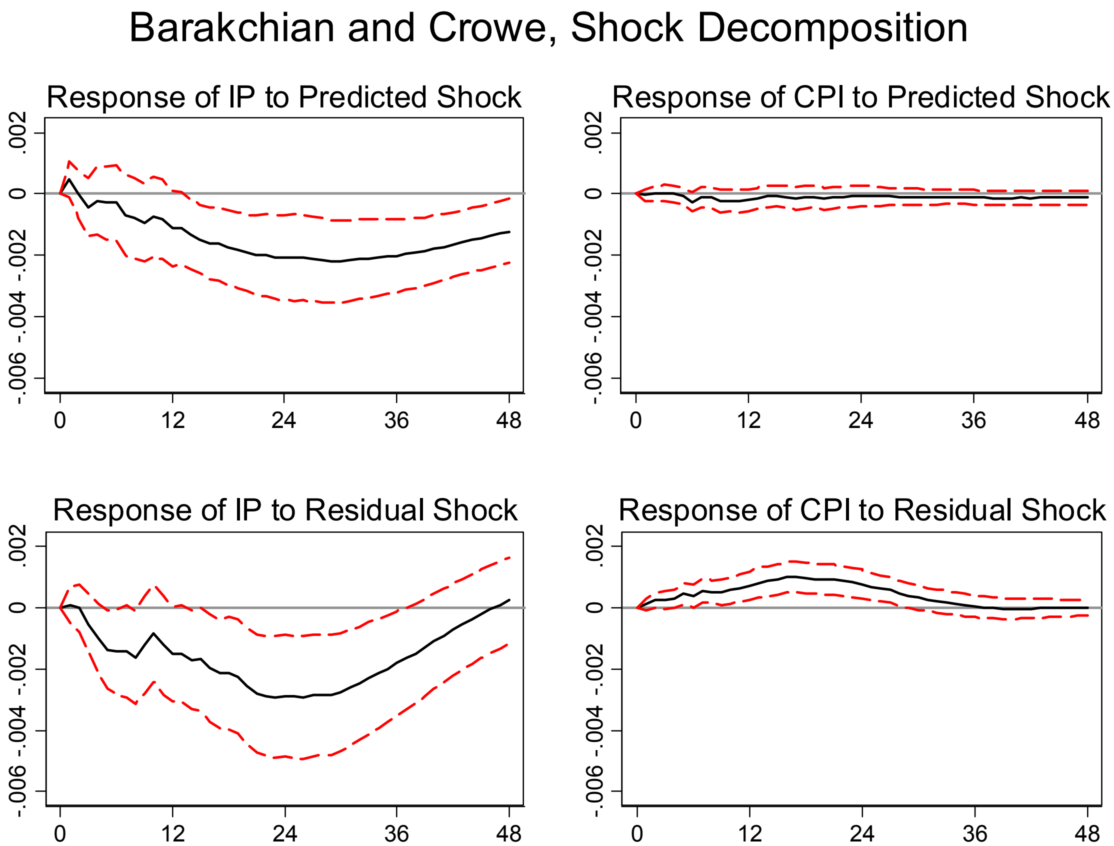
## VI. CONCLUSION

Conventional VAR and non-VAR identification schemes for estimating the effect of U.S. monetary policy shocks on the wider economy are sensitive to the sample period under consideration. In particular, these schemes generate unrealistic impulse response functions for output, and to a lesser extent prices, for the period since the mid-1980s known as the great moderation. We show that these apparently perverse results may be generated by a relative decline in the shock component of policy (making the effect of shocks harder to identify), and a failure to properly identify the Fed's reaction function to allow for changes in its parameters over time, particularly a greater weight placed on forward-looking variables.

We outline a new measure of monetary policy shocks derived from Fed Funds futures contracts that is less prone to these problems. As a result, our new measure generates a more realistic impulse response function for output, with a small but statistically significant negative effect whose maximum impact is felt at a horizon of two years following a monetary contraction. We also find evidence for a price puzzle over the medium term. We find that almost half of output variability (at a 3 year horizon) can be explained by monetary policy shocks using our new identification strategy, twice the share under other identification schemes for the same period. While our shock measure may be contaminated by the Fed's systematic policy reaction to its private forecasts, this is likely to bias our estimated impulse responses towards zero, so that the estimated output response may represent an underestimate. Moreover, while this simultaneity bias appears to be small under our identification scheme, it is likely to be more important for VAR-based identification methods.

We would rationalize the high share of output volatility accounted for by our shock measure by a combination of substantive and econometric factors. Substantively, the Fed exercised more effective control over the economy during the 'great moderation' period covered in our analysis, partly via an improved focus on forward-looking indicators, helping to minimize the impact of exogenous demand shocks so that a greater share of the remaining shocks is accounted for by policy itself. Although the absolute effect of the shocks is small, their relative impact is large in a

**Figure 9. Impulse Response Functions for Decomposed Shock Measure**




Structural VAR (Monthly data, 4 endogenous variables plus constant and linear time trend, 12 lags). Data sample 1988:12-2002:12.
Variables ordered as industrial production, consumer prices, both seasonally adjusted and in logs, and the predicted and residual components of the regression of our shock measure on the Fed's private information described in the text, both cumulated. Graphs show response of industrial production and CPI to a one standard deviation positive shock to each policy measure.
Structural shocks obtained via Cholesky decomposition.
Two Standard Error bands produced by parametric bootstrapping (500 replications).

period of relatively low overall volatility. In addition, our shock measure captures only policy changes that were truly unanticipated by the private sector, and it is these unexpected monetary policy changes that are generally believed to have the largest impact on output. Additional econometric factors include the fact that our shock variable is not a pure measure of shocks but also includes the Fed's systematic response to its private forecasts. While the inclusion of the Fed's response to private information will tend to reduce the magnitude of the estimated coefficients, it may increase the overall effect by increasing the size of the estimated shocks, although we find that this systematic component accounts for only one-third of our shock measure's total estimated contribution to output variation.

Our results have a number of implications for further research. For instance, we hope to use our new measure to assess the impact of U.S. monetary policy shocks on other economies, contributing to the literature on international monetary policy spillovers.

## APPENDIX

### Existing Identification Schemes: Details and Estimation

**Christiano, Eichenbaum and Evans (1996)** estimate a quarterly VAR with six variables and four lags over the period 1960Q1-1992Q4. They employ a recursive VAR with the variables ordered as: the log of real GDP, the log of the GDP deflator, the log of a commodity price index, the log of nonborrowed reserves (NBR), the federal funds rate, and the log of total reserves. They use two different proxies for monetary shocks: the residual in the Federal Funds rate equation and the residual in the nonborrowed reserves equation.[42] We focus on the former as it seems most relevant for the recent period—when the Federal Funds rate is clearly the monetary instrument—and in any case their baseline results are very similar across both specifications. Our VAR also includes a constant and a linear time trend. All series except for the Fed Funds rate are seasonally adjusted.

**Bernanke and Mihov (1998)** develop a model in which the relationships among macroeconomic variables are left unrestricted while contemporaneous identification restrictions on monetary variables which are relevant to the market for commercial bank reserves (total, borrowed and non-borrowed reserves, respectively) are imposed in order to model the Fed's operating procedure. They consider four different monetary regimes: Federal Funds rate targeting, nonborrowed reserves targeting, orthogonalized nonborrowed reserves targeting, and borrowed reserves targeting, where all four models are over-identified, and conclude that the Fed's procedures have changed over time so that no single model is optimal for the whole 1965-1996 period. In particular, they find strong evidence that the Fed switched to nonborrowed reserves targeting during the 1979-1982 period, but that the Federal Funds rate targeting model does well for the pre-1979 period and exceptionally well for the post-1988 period. We therefore focus on this latter model. The variables in our VAR are ordered as log industrial production, log consumer price index, log commodity prices, the fed funds rate and total reserves and non-borrowed reserves, all variables except for the Fed Funds rate are seasonally adjusted and the last two variables are normalized by dividing by the 36-month moving average of total reserves. The VAR is estimated with 13 lags and includes a constant and a linear time trend. Consistent with the restrictions implied by the Federal Funds rate targeting regime, as in Bagliano and Favero (1998), one over-identification restriction is imposed on the contemporaneous coefficient matrix.[43]

---

[42]When nonborrowed reserves is the instrument, the ordering of the federal funds rate and nonborrowed reserves is reversed.

[43]Hence, the matrix of contemporanous coefficients $A_0$ is lower triangular (there is a single over-identifying restriction, imposing a value of $-1$ on the penultimate element of the last row of $A_0$ ). Bernanke and Mihov (1998) suggest that a more general estimation method is to use a just-identified model to measure the monetary policy shock such that this policy shock is most consistent with the estimates of the Federal Reserve's operating procedure in each period, rather than choosing a specific monetary regime for the whole period.

**Sims and Zha (2006)** include a somewhat different list of variables from most other studies, including the producers' price index components for crude materials and intermediate materials and a measure of bankruptcies.[44] They consider two sets of different variables to represent monetary policy: total reserves and the Federal Funds rate on the one hand, and M2 and the Treasury Bill rate on the other. We focus on the M2-Treasury Bill rate model, because the response of output to a contractionary monetary policy shock is more significantly negative in the original period in this model, helping to sharpen the contrast with the recent period. Their quarterly model, which is estimated over the period 1964Q1-1994Q4, includes the 90-day treasury bill rate, M2, the GNP deflator, real GNP, average hourly earnings of nonagricultural workers, the producers' price index for intermediate materials, bankruptcy filings, and the producers' price index for crude materials. Their set of identifying assumptions allows them to identify their nonrecursive VAR.[45]

**Romer and Romer (2004)** argue that these means of identifying monetary policy shocks are subject to two major deficiencies. First, monetary policy instruments respond endogenously to other economic variables. For instance, the effective Federal Funds rate may differ from the target rate due to movements in the reserves market. This will lead to biased estimates of monetary policy if the endogeneity is not corrected for. Second, monetary instruments are not adjusted at random, but respond to policymakers' beliefs about the future path of the economy. This will introduce a further source of bias. In order to overcome the problem of endogeneity, Romer and Romer (2004) derive a series of intended Fed Funds rate changes by combining information on the effective Funds rate with the Fed's narrative accounts of each FOMC meeting. Then, in order to remove anticipatory movements from the series, they regress the intended funds rate on the Fed's internal (Greenbook) forecasts of inflation and real activity. The residual from this regression, i.e. the change in the intended funds rate which is not a reaction to the future economic conditions, is used as a proxy for the underlying monetary policy shock $s_t$. This new series is constructed for the period 1969-1996.

Romer and Romer estimate a monthly VAR with three variables including industrial production, PPI for finished goods and their measure of the monetary policy shock. Using Christiano and others' recursive identification strategy, they assume that monetary policy responds contemporaneously to other variables but does not affect them contemporaneously.

---

[44]They argue that the crude materials price index is a rapidly responding, quickly observable, indicator of possible inflationary pressure to which monetary policy might respond rapidly, and that its inclusion therefore helps to correctly estimate the monetary policy reaction function, while the intermediate materials price index and bankruptcies are important in the transmission of monetary policy to the economy, and so their inclusion helps to correctly estimate the response of the private sector to monetary policy shocks. In fact, the contribution of the bankruptcy variable to the results is only modest.

[45]See Sims and Zha (2006), Table 3, for the full list of the identifying assumptions used in our estimation (the last row and column are redundant in our case as we drop the bankruptcies variable, which is ordered last). Their non-recursive identification restrictions include $\frac{n(n-1)}{2}$ zero restrictions on $\mathbf{A}_0$, $\frac{n(n-1)}{2}$ orthogonality restrictions and $n$ normalization restrictions on $\mathbf{\Psi}$ and two over-identifying restrictions on the money demand equation (the second row of $\mathbf{A}_0$). All variables with the exception of the Treasury Bill rate are seasonally adjusted and enter in logs. The VAR includes four lags of the endogenous variables and a constant and linear time trend.

## Data Sources and Construction

### New shock measure

Fed Funds Futures prices are obtained on an end-of-day basis from Bloomberg (generic Fed Funds futures contracts with Bloomberg mnemonics *FF1-FF6* for the current month through 5 months ahead). Non-trading days take on the value of the last previous trading day. Data on the volume of contracts is available from Bloomberg from July 1989, and suggests that trading volumes were low in the early months. Even as late as 1993 some FOMC meeting days saw no trading in the current month futures contract (as late as 2001 for the 5 months ahead contract). Since there is always trading in contracts at two or more maturities on meeting days, even when not all contracts trade, there is a clear benefit in combining information from contracts at various maturities, rather than relying on a particular maturity. The volume of trading has grown markedly over time. To illustrate this, table A1 compares (average) monthly means and standard deviations of trading volumes (number of contracts) at different maturities for 1997 and 2007.

**Table A1. Average monthly SD and mean for trading volume, Futures contracts at various horizons**

| | 1997 | | 2007 | |
|---|---|---|---|---|
| Horizon (+ months) | Mean | SD | Mean | SD |
| 0 | 1,216 | 1,048 | 9,677 | 9,582 |
| 1 | 1,926 | 1,424 | 17,060 | 12,245 |
| 2 | 1,226 | 980 | 13,383 | 8,498 |
| 3 | 522 | 449 | 10,917 | 6,546 |
| 4 | 185 | 190 | 7,953 | 5,153 |
| 5 | 50 | 57 | 3,979 | 2,838 |

Annual average of monthly SD and mean for trading volume (number of contracts traded), for Fed Funds futures contracts at different horizons. Authors' estimates based on data from Bloomberg.

Following Kuttner (2001), the impact of policy announcements (or non-announcements) following FOMC meetings is estimated by comparing the end of day price on the day following the (last) day of the meeting with that on the meeting day for meetings occurring before February 1994, and comparing the price on the day of the meeting with that on the day before the meeting for subsequent meetings. Like Kuttner (2001), the spot month shock is replaced by the one month ahead shock for meetings occurring during the last 3 days of the month. Again following Kuttner (2001), appropriate adjustments to the formulae are also made for shocks during the last day of the month (pre-1994) and first day of the month (post-1994).

Table A2 presents the results of regressing the Fed's private information on the forecast error associated with the public information, to give an indication of the degree to which the Fed's private information set represents a real information advantage (see the discussion in section V of the paper).

**Table A2. Fed Private Information**

| Variable (Fed's Private Information) | GDP | | | GDP Deflator | | |
|---|---|---|---|---|---|---|
| Horizon | Q | Q+1 | Q+2 | Q | Q+1 | Q+2 |
| Private Sector Forecast Error | -0.156*** | -0.085** | -0.080*** | -0.373*** | -0.112** | -0.079* |
| | (0.043) | (0.034) | (0.030) | (0.068) | (0.046) | (0.041) |
| Constant | -0.056 | -0.064 | -0.007 | -0.005 | -0.127*** | -0.202*** |
| | (0.041) | (0.076) | (0.067) | (0.053) | (0.040) | (0.036) |
| Observations | 120 | 120 | 120 | 120 | 120 | 120 |
| R-Squared | 0.01 | 0.05 | 0.06 | 0.20 | 0.05 | 0.03 |

The dependent variable in each regression is the Fed's private information, given by the difference between the Fed Greenbook and Blue Chip forecasts (as a proxy for private sector expectations); the right hand side variable is the private sector's forecast error, given by the difference between the Blue Chip forecast and the actual outturn. A negative coefficient on this variable indicates that the Fed's private information is closer to the actual than the private sector's information (i.e., the Fed has useful private information); the $R^2$ indicates the share of the Fed's private information that is accurate, ex post. Standard Errors in parentheses (significance levels denoted as ***: 1 percent; **: 5 percent; *: 10 percent). 120 per-meeting observations (February 1988-December 2002).

**Romer and Romer Shock Measure**

The series was extended as far as possible using the methodology of Romer and Romer (2004). As a first stage, the intended Federal Funds rate was obtained by a reading of the minutes and statements released following the relevant FOMC meeting. These were obtained from the Federal Reserve Board of Governors website, at http://www.federalreserve.gov/monetarypolicy/fomccalendars.htm (2003 onwards) and http://www.federalreserve.gov/monetarypolicy/fomc_historical.htm (1997 through 2002). In all but three cases, the desired policy change (*DTARG*) is equal to the actual change in rates (Table A3).

To arrive at the measure of policy shocks, *DTARG* is then regressed on a number of forecast variables from the Greenbook forecasts. These variables include the GDP deflator, real GDP and unemployment. These forecasts were obtained for 1997-2002 from the Federal Reserve Board of Governors website, http://www.federalreserve.gov/monetarypolicy/fomc_historical.htm.
The GDP deflator growth rate (in percentage points) is calculated as:

$$\Delta Deflator = 100 \times \left( \frac{100 + \Delta \text{ Nominal GDP}}{100 + \Delta \text{ Real GDP}} - 1 \right)$$

Real GDP growth and the unemployment rate were simply taken from the published forecasts. To match the forecast quarters to the meeting, the current quarter (quarter 0) is taken as the quarter that the meeting took place in (the concluding day for two day meetings).
Since Greenbook forecasts are not currently available for 2003 onwards, we proxied for the Greenbook forecasts using the consensus forecast available from the Blue Chip Economic Indicators. Although it has been established that the Greenbook Forecasts have a lower forecast error than private sector forecasts, the difference in terms of mean square error is not huge for the Blue Chip forecasts (Romer and Romer, 2000). In addition, the Blue Chip forecast format

(monthly forecasts for quarterly horizons) follows closely that of the Greenbook, enabling them to be mapped more easily into the existing framework than other commercial forecast series (e.g. those of Consensus Economics or the Philadelphia Fed's Survey of Professional Forecasters). Nevertheless, using the Blue Chip forecasts represent a second-best response to the non-availability of the Greenbook forecasts over this period, as we are really interested in the information set available to the FOMC policymakers, which will contain elements missing from the Blue Chip forecasts.

In terms of matching the Blue Chip forecasts to the FOMC meetings, the Blue Chip forecasts are released on the 10th of each month; but in reality they are circulated a couple of days before this, and the data collection likely predates this by several days. We assume that the data collection is complete by the 5th of each month. For meetings before the 5th of each month, we use the forecasts from the previous month; for meetings on or after the 5th, we use the current month's forecasts. Since the Blue Chip forecasts do not include a nominal GDP forecast, we use the change in the GDP chained price index as a proxy for the change in the GDP deflator.

Our extended Romer and Romer shock measure is then calculated using the OLS residuals from the policy rule for the entire period. Like Romer and Romer, monthly shocks are generated as the sum of per-meeting shocks during the month in question, with zero for months where no meeting took place. A cumulative monthly measure is then entered in the VAR.

We combine this information and the data available from the Romer and Romer (2004) data appendix to produce three separate series:

$RESID96$ : the residual using only the original Romer and Romer data. This is identical to their variable RESID, and is calculated only to ensure that we had accurately replicated their methodology. It is available up to end-1996.

$RESID02$ : the residual using the original Romer and Romer data and our new data up to end-2002. This series therefore only covers the period for which we have Greenbook forecasts.

$RESID08$ : the residual using the original Romer and Romer data and all our new data, including the Blue Chip forecasts from start-2003 up to June 2008.

The regression equations from which these residual series are derived are given in Table A4. Note that the coefficient estimates are more or less unchanged by the increase in the sample, even though the full sample represents a 35 percent increase, in terms of meeting observations, compared to the original sample.

We then use Romer and Romer's methodology to convert these by-meetings data into monthly shock and monthly cumulative shock series.

The data used in this paper, excluding the Blue Chip forecasts (which are proprietary), as well as our shock series and the updated Romer and Romer shock series are available online at http://www.imf.org/external/pubs/ft/wp/2010/data/wp10230.zip.

**Macroeconomic data**

The commodity price index was obtained from the IMF's Commodities unit. All other macroeconomic data was obtained from the Federal Reserve Bank of St. Louis (*FRED* database).

**Table A3. Narrative of FOMC Meetings**

| Meeting Date | OLDTARG | DTARG | Narrative | Source |
|---|---|---|---|---|
| 2/5/1997 | 5.25 | 0 | No Change, slight bias toward tightening, one dissent | FOMC Minutes |
| 3/25/1997 | 5.25 | 0.25 | Up 25 bps, symmetric bias, no dissent | FOMC Minutes |
| 5/20/1997 | 5.5 | 0 | No Change, slight bias toward tightening, one dissent | FOMC Minutes |
| 7/2/1997 | 5.5 | 0 | No change, symmetric bias, no dissent | FOMC Minutes |
| 8/19/1997 | 5.5 | 0 | No change, slight bias toward tightening, no dissent | FOMC Minutes |
| 9/30/1997 | 5.5 | 0 | No change, slight bias toward tightening, no dissent | FOMC Minutes |
| 11/12/1997 | 5.5 | 0 | No change, slight bias toward tightening, one dissent | FOMC Minutes |
| 12/16/1997 | 5.5 | 0 | No change, symmetric bias, one dissent | FOMC Minutes |
| 2/4/1998 | 5.5 | 0 | No change, symmetric bias, no dissent | FOMC Minutes |
| 3/31/1998 | 5.5 | 0 | No change, slight bias toward tightening, one dissent | FOMC Minutes |
| 5/19/1998 | 5.5 | 0 | No change, slight bias toward tightening, two dissents | FOMC Minutes |
| 7/1/1998 | 5.5 | 0 | No change, slight bias toward tightening, one dissent | FOMC Minutes |
| 8/18/1998 | 5.5 | 0 | No change, symmetric bias, one dissent | FOMC Minutes |
| 9/29/1998 | 5.5 | -0.25 | Down 25bps, slight bias toward loosening, no dissent | FOMC Minutes |
| 11/17/1998 | 5 | -0.25 | Reduced 25bps at Oct 15 Conference call; further 25bps reduction. Symmetric bias, one dissent | FOMC Minutes |
| 12/22/1998 | 4.75 | 0 | No change, symmetric bias, no dissent | FOMC Minutes |
| 2/3/1999 | 4.75 | 0 | No change, symmetric bias, no dissent | FOMC Minutes |
| 3/30/1999 | 4.75 | 0 | No change, symmetric bias, no dissent | FOMC Minutes |
| 5/18/1999 | 4.75 | 0.125 | No change, but strong bias toward tightening including press statement, no dissent | FOMC Minutes + Statement |
| 6/30/1999 | 4.75 | 0.25 | Up 25 bps, symmetric bias, one dissent | FOMC Minutes |
| 8/24/1999 | 5 | 0.25 | Up 25 bps, symmetric bias, one dissent | FOMC Minutes |
| 10/5/1999 | 5.25 | 0 | No change, asymmetric bias but specifically signalled did not commit FOMC to increase in short term. No dissent | FOMC Minutes + Statement |
| 11/16/1999 | 5.25 | 0.25 | Up 25 bps, symmetric bias, no dissent | FOMC Minutes |
| 12/21/1999 | 5.5 | 0 | No change, some concern about inflationary risks but symmetric bias, no dissent. | FOMC Minutes + Statement |

**Table A3. Narrative of FOMC Meetings**

| Meeting Date | OLDTARG | DTARG | Narrative | Source |
|---|---|---|---|---|
| 2/2/2000 | 5.5 | 0.25 | Up 25 bps, bias toward tightening, no dissent | FOMC Minutes |
| 3/21/2000 | 5.75 | 0.25 | Up 25 bps, bias toward tightening, no dissent | FOMC Minutes |
| 5/16/2000 | 6 | 0.5 | Up 50 bps, bias toward tightening, no dissent | FOMC Minutes |
| 6/28/2000 | 6.5 | 0 | No change, bias toward tightening, no dissent | FOMC Minutes |
| 8/22/2000 | 6.5 | 0 | No change, bias toward tightening, no dissent | FOMC Minutes |
| 10/3/2000 | 6.5 | 0 | No change, bias toward tightening, no dissent | FOMC Minutes |
| 11/15/2000 | 6.5 | 0 | No change, bias toward tightening, no dissent | FOMC Minutes |
| 12/19/2000 | 6.5 | -0.125 | No change, but shift in bias toward loosening and discussion stressed need to move in intervening weeks if necessary. Support for immediate cut but no official dissent. Cut to 6 at conference call on Jan 03, 2001. | FOMC Minutes |
| 1/25/2001 | 6 | -0.5 | Down 50 bps, bias toward loosening, no dissent | FOMC Minutes |
| 3/20/2001 | 5.5 | -0.5 | Down 50 bps, bias toward loosening, no dissent. Further conference calls on April 11 and 18, rate cut 50bps at latter meeting, no dissent. | FOMC Minutes |
| 5/15/2001 | 4.5 | -0.5 | Down 50 bps, bias toward loosening, one dissent | FOMC Minutes |
| 6/27/2001 | 4 | -0.25 | Down 25bps, bias toward loosening, one dissent | FOMC Minutes |
| 8/21/2001 | 3.75 | -0.25 | Down 25bps, bias toward loosening, no dissent. Subsequent meetings on September 13 and 17; rate cut 50bps at latter meeting, no dissent. | FOMC Minutes |
| 10/2/2001 | 3 | -0.5 | Down 50bps, bias toward loosening, no dissent. | FOMC Minutes |
| 11/6/2001 | 2.5 | -0.5 | Down 50bps, bias toward loosening, no dissent. | FOMC Minutes |
| 12/11/2001 | 2 | -0.25 | Down 25bps, bias toward loosening, one dissent | FOMC Minutes |
| 1/30/2002 | 1.75 | 0 | No change, bias toward loosening, no dissent | FOMC Minutes |
| 3/19/2002 | 1.75 | 0 | No change, symmetric balance of risks, no dissent | FOMC Minutes |
| 5/7/2002 | 1.75 | 0 | No change, symmetric balance of risks, no dissent | FOMC Minutes |
| 6/26/2002 | 1.75 | 0 | No change, symmetric balance of risks, no dissent | FOMC Minutes |

## Table A3. Narrative of FOMC Meetings

| Meeting Date | OLDTARG | DTARG | Narrative | Source |
|---|---|---|---|---|
| 8/13/2002 | 1.75 | 0 | No change, balance of risk tilted toward economic weakness but minutes reveal no desire to move quickly on rates. No dissent. | FOMC Minutes |
| 9/24/2002 | 1.75 | -0.125 | No change, balance of risk tilted toward economic weakness. Minutes reveal some desire to move in inter-meeting period if subsequent information, and also some desire to move at meeting. Two dissents. | FOMC Minutes |
| 11/6/2002 | 1.75 | -0.5 | 50 bps reduction, but symmetric balance of risks. No dissent. | FOMC Minutes |
| 12/10/2002 | 1.25 | 0 | No change, symmetric balance of risks, no dissent | FOMC Minutes |
| 1/29/2003 | 1.25 | 0 | No change, symmetric balance of risks, no dissent | FOMC Minutes |
| 3/18/2003 | 1.25 | 0 | No change, symmetric balance of risks, no dissent | FOMC Minutes |
| 5/6/2003 | 1.25 | 0 | No change, balance of risks toward weakness, no dissent | FOMC Minutes + Statement |
| 6/25/2003 | 1.25 | -0.25 | 25 bps reduction, balance of risk toward downside, one dissent | FOMC Minutes |
| 8/12/2003 | 1 | 0 | No change, balance of risks toward weakness, no dissent | FOMC Minutes |
| 9/16/2003 | 1 | 0 | No change, balance of risks toward weakness, no dissent | FOMC Minutes |
| 10/28/2003 | 1 | 0 | No change, balance of risks toward weakness, no dissent | FOMC Minutes |
| 12/9/2003 | 1 | 0 | No change, symmetric balance of risks, no dissent | FOMC Minutes |
| 1/28/2004 | 1 | 0 | No change, symmetric balance of risks, no dissent | FOMC Minutes |
| 3/16/2004 | 1 | 0 | No change, symmetric balance of risks, no dissent | FOMC Minutes |
| 5/4/2004 | 1 | 0 | No change, symmetric balance of risks, no dissent | FOMC Minutes |
| 6/30/2004 | 1 | 0.25 | 25 bps increase, symmetric balance of risks, no dissent | FOMC Minutes |
| 8/10/2004 | 1.25 | 0.25 | 25 bps increase, symmetric balance of risks, no dissent | FOMC Minutes |
| 9/21/2004 | 1.5 | 0.25 | 25 bps increase, symmetric balance of risks, no dissent | FOMC Minutes |
| 11/10/2004 | 1.75 | 0.25 | 25 bps increase, symmetric balance of risks, no dissent | FOMC Minutes |
| 12/14/2004 | 2 | 0.25 | 25 bps increase, symmetric balance of risks, no dissent | FOMC Minutes |
| 2/2/2005 | 2.25 | 0.25 | 25 bps increase, symmetric balance of risks, no dissent | FOMC Minutes |

**Table A3. Narrative of FOMC Meetings**

| Meeting Date | OLDTARG | DTARG | Narrative | Source |
|---|---|---|---|---|
| 3/22/2005 | 2.5 | 0.25 | 25 bps increase, symmetric balance of risks, no dissent | FOMC Minutes |
| 5/3/2005 | 2.75 | 0.25 | 25 bps increase, symmetric balance of risks, no dissent | FOMC Minutes |
| 6/30/2005 | 3 | 0.25 | 25 bps increase, symmetric balance of risks, no dissent | FOMC Minutes |
| 8/9/2005 | 3.25 | 0.25 | 25 bps increase, symmetric balance of risks, no dissent | FOMC Minutes |
| 9/20/2005 | 3.5 | 0.25 | 25 bps increase, symmetric balance of risks, one dissent | FOMC Minutes |
| 11/1/2005 | 3.75 | 0.25 | 25 bps increase, symmetric balance of risks, no dissent | FOMC Minutes |
| 12/13/2005 | 4 | 0.25 | 25 bps increase, signal that some additional policy firming may be necessary, no dissent | FOMC Minutes |
| 1/31/2006 | 4.25 | 0.25 | 25 bps increase, signal that some additional policy firming may be necessary, no dissent | FOMC Minutes |
| 3/28/2006 | 4.5 | 0.25 | 25 bps increase, signal that some additional policy firming may be necessary, no dissent | FOMC Minutes |
| 5/10/2006 | 4.75 | 0.25 | 25 bps increase, signal that additional policy firming will depend on new information, no dissent | FOMC Minutes |
| 6/29/2006 | 5 | 0.25 | 25 bps increase, signal that additional policy firming will depend on new information, no dissent | FOMC Minutes |
| 8/8/2006 | 5.25 | 0 | No change, signal that inflation risks predominate and might require further tightening, one dissent | FOMC Minutes |
| 9/20/2006 | 5.25 | 0 | No change, signal that inflation risks predominate and might require further tightening, one dissent | FOMC Minutes |
| 10/25/2006 | 5.25 | 0 | No change, signal that inflation risks predominate and might require further tightening, one dissent | FOMC Minutes |
| 12/12/2006 | 5.25 | 0 | No change, signal that inflation risks predominate and might require further tightening, one dissent | FOMC Minutes |
| 1/31/2007 | 5.25 | 0 | No change, signal that inflation risks predominate and might require further tightening, no dissent | FOMC Minutes |
| 3/21/2007 | 5.25 | 0 | No change, signal that inflation risks predominate and might require further tightening, no dissent | FOMC Minutes |

**Table A3. Narrative of FOMC Meetings**

| Meeting Date | OLDTARG | DTARG | Narrative | Source |
|---|---|---|---|---|
| 5/9/2007 | 5.25 | 0 | No change, signal that inflation risks predominate and might require further tightening, no dissent | FOMC Minutes |
| 6/28/2007 | 5.25 | 0 | No change, signal that inflation risks predominate and might require further tightening, no dissent | FOMC Minutes |
| 8/7/2007 | 5.25 | 0 | No change, signal that downside risks growing but inflation risk predominates, no dissent. Subsequent conference calls on August 10 and 16 but no policy action. | FOMC Minutes |
| 9/18/2007 | 5.25 | -0.5 | 50 bps reduction, statement speaks of increased economic uncertainty, no dissent. | FOMC Minutes |
| 10/31/2007 | 4.75 | -0.25 | 25 bps reduction, risks judged symmetric following policy change, one dissent. | FOMC Minutes |
| 12/11/2007 | 4.5 | -0.25 | 25 bps reduction, statement speaks of increased uncertainty over growth and inflation, one dissent. Subsequent conference calls on Jan. 9 and 21; rate reduced 75bps at latter meeting. | FOMC Minutes |
| 1/30/2008 | 3.5 | -0.5 | 50 bps reduction, statement that downside risks remain. One dissent. Conference call on March 10 to discuss financial market developments. | FOMC Minutes |
| 3/18/2008 | 3 | -0.75 | 75 bps reduction, statement mentions downside risks to growth. Two dissents. | FOMC Minutes |
| 4/30/2008 | 2.25 | -0.25 | 25 bps reduction, statement mentions inflation concerns. Two dissents. | FOMC Minutes |
| 6/25/2008 | 2 | 0 | No change, statement mentions inflation concerns. One dissent. | FOMC Minutes |

Table shows updates to Romer and Romer's (2004) variables *OLDTARG* (Fed Funds target rate going into meeting) and *DTARG* (desired change in target rate decided at meeting), based on authors' reading of FOMC Minutes and Statements.

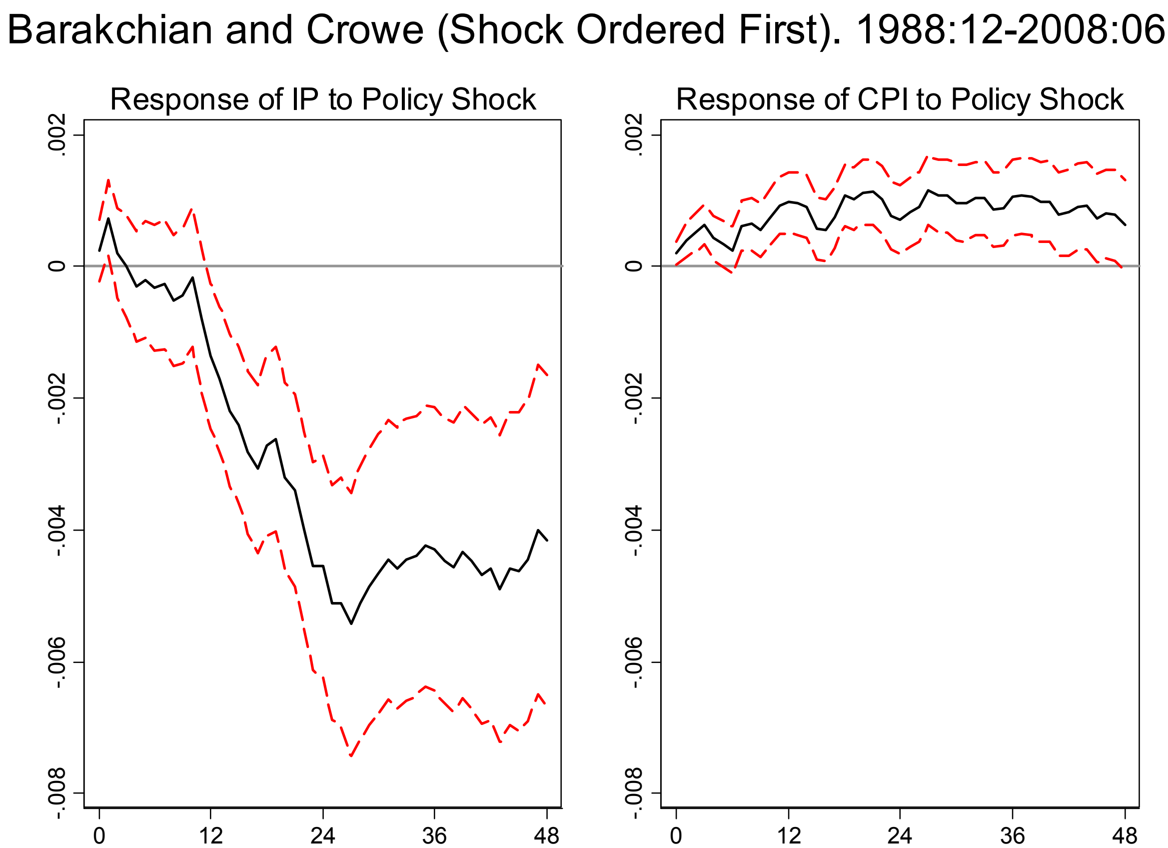
**Table A4. ΔFed Funds regressions to obtain residuals (Romer and Romer shocks)**

| | (1) Original Sample | (2) Greenbook Only | (3) Full Sample |
|---|---|---|---|
| Fed $\text{Funds}_{-1}$ | -0.021* | -0.018* | -0.018* |
| | (0.012) | (0.011) | (0.009) |
| Output $\text{Growth}_{-1}$ | 0.007 | 0.007 | 0.008 |
| | (0.010) | (0.009) | (0.009) |
| Output $\text{Growth}_0$ | 0.003 | 0.007 | 0.009 |
| | (0.019) | (0.017) | (0.016) |
| Output $\text{Growth}_1$ | 0.01 | 0.02 | 0.02 |
| | (0.032) | (0.025) | (0.024) |
| Output $\text{Growth}_2$ | 0.022 | 0.013 | 0.012 |
| | (0.032) | (0.025) | (0.024) |
| GDP $\text{Deflator}_{-1}$ | 0.021 | 0.027 | 0.028 |
| | (0.024) | (0.021) | (0.020) |
| GDP $\text{Deflator}_0$ | -0.044 | -0.043 | -0.041 |
| | (0.029) | (0.026) | (0.025) |
| GDP $\text{Deflator}_1$ | 0.01 | 0.008 | 0.01 |
| | (0.044) | (0.040) | (0.038) |
| GDP $\text{Deflator}_2$ | 0.052 | 0.053 | 0.048 |
| | (0.047) | (0.043) | (0.041) |
| $\text{Unemployment}_0$ | -0.048** | -0.045** | -0.045*** |
| | (0.021) | (0.018) | (0.017) |
| ∆Output $\text{Growth}_{-1}$ | 0.050* | 0.048* | 0.040* |
| | (0.030) | (0.026) | (0.024) |
| ∆Output $\text{Growth}_0$ | 0.152*** | 0.138*** | 0.138*** |
| | (0.030) | (0.026) | (0.025) |
| ∆Output $\text{Growth}_1$ | 0.021 | 0.016 | 0.02 |
| | (0.046) | (0.038) | (0.036) |
| ∆Output $\text{Growth}_2$ | 0.021 | 0.03 | 0.029 |
| | (0.051) | (0.044) | (0.042) |
| ∆GDP $\text{Deflator}_{-1}$ | 0.057 | 0.054 | 0.055 |
| | (0.045) | (0.039) | (0.036) |
| ∆GDP $\text{Deflator}_0$ | 0.003 | -0.01 | -0.011 |
| | (0.048) | (0.043) | (0.040) |
| ∆GDP $\text{Deflator}_1$ | 0.031 | 0.034 | 0.034 |
| | (0.074) | (0.066) | (0.063) |
| ∆GDP $\text{Deflator}_2$ | -0.062 | -0.061 | -0.064 |
| | (0.081) | (0.074) | (0.070) |
| Constant | 0.171 | 0.08 | 0.077 |
| | (0.141) | (0.109) | (0.096) |
| Observations | 263 | 311 | 355 |
| R-squared | 0.282 | 0.284 | 0.286 |

OLS regression results, from regression of change in desired change in Fed Funds rate on the rate going into the meeting (Fed $\text{Funds}_{-1}$) and 17 Greenbook variables, as outlined in Romer and Romer (2004). Results in the first column show results for the original Romer and Romer sample; the second column includes data to 2002 (i.e. only Greenbook data); the third column includes data to June 2008 (substituting Blue Chip data for Greenbook data for 2003-08, as described in the paper). The residual from the first column is used as the Romer and Romer shock for replicating their results (Figure 4, panel a); the residual from the third column is used as the Romer and Romer shock for the post-1988 period (Figure 4, panel b). Standard Errors in parentheses (significance levels denoted as ***: 1 percent; **: 5 percent; *: 10 percent).

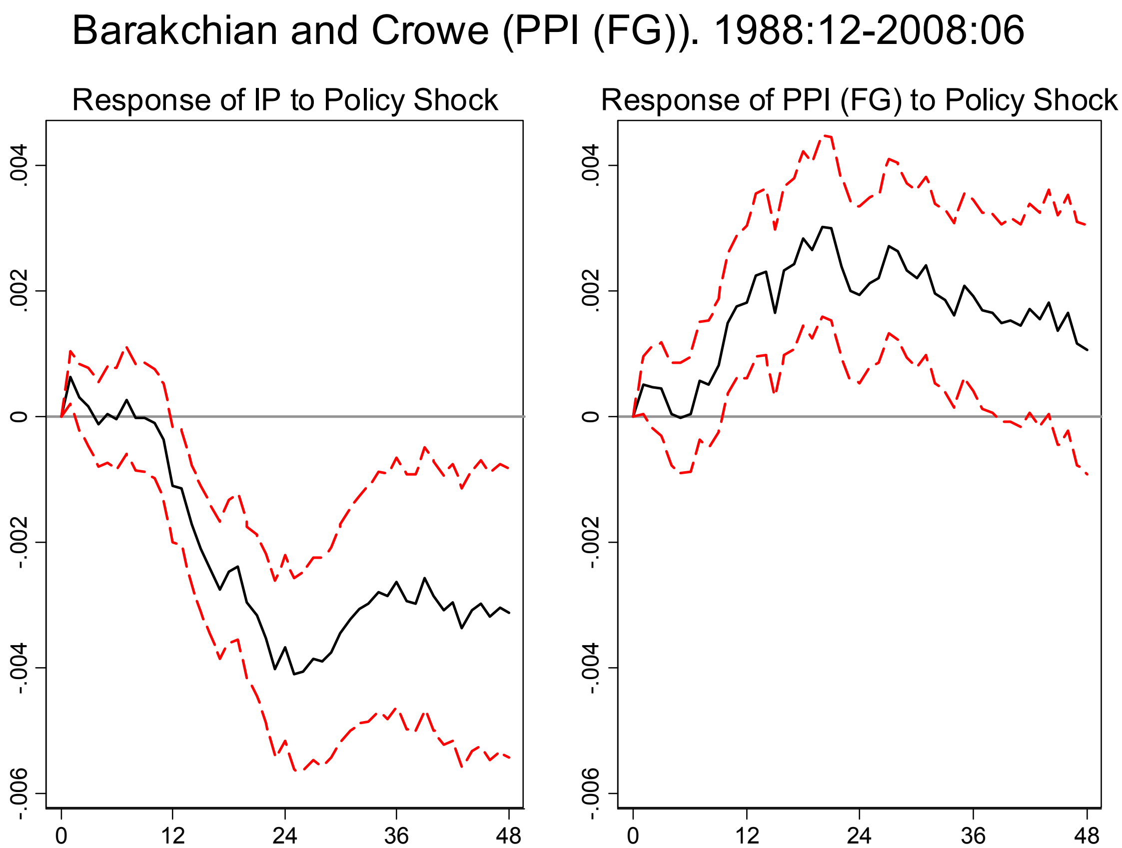
Additional Figures (Robustness Checks).

**Figure A.1 Shock Ordered First**




Structural VAR (Monthly data, 3 endogenous variables plus constant and linear time trend, 36 lags).
Variables ordered as our shock measure, cumulated; industrial production, consumer prices, both seasonally adjusted and in logs. Graphs show response of industrial production and CPI to a one standard deviation positive shock to the policy measure.
Structural shocks obtained via Cholesky decomposition.
Two Standard Error bands produced by parametric bootstrapping.

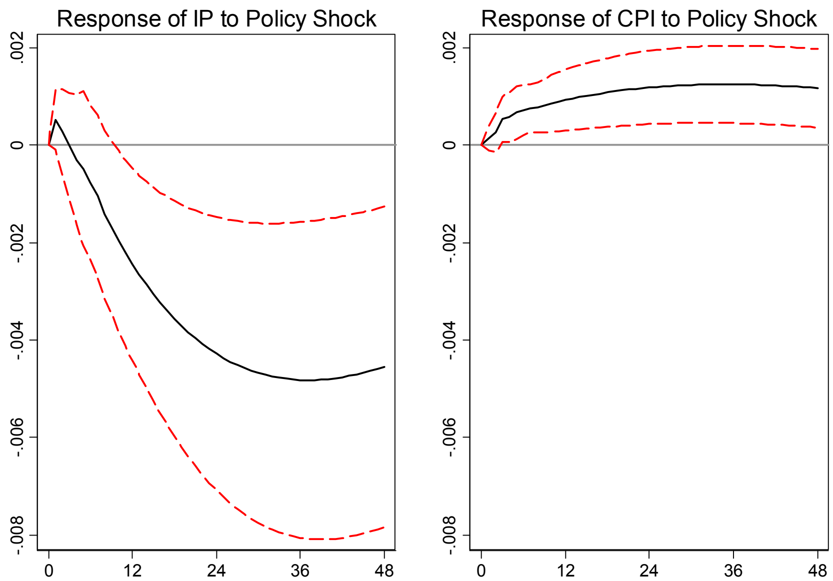
**Figure A2. PPI (Finished Goods) as Price Measure**




Structural VAR (Monthly data, 3 endogenous variables plus constant and linear time trend, 36 lags).
Variables ordered as industrial production, PPI (finished goods), both seasonally adjusted and in logs, and our shock measure, cumulated. Graphs show response of industrial production and PPI (FG) to a one standard deviation positive shock to the policy measure.
Structural shocks obtained via Cholesky decomposition.
Two Standard Error bands produced by parametric bootstrapping.

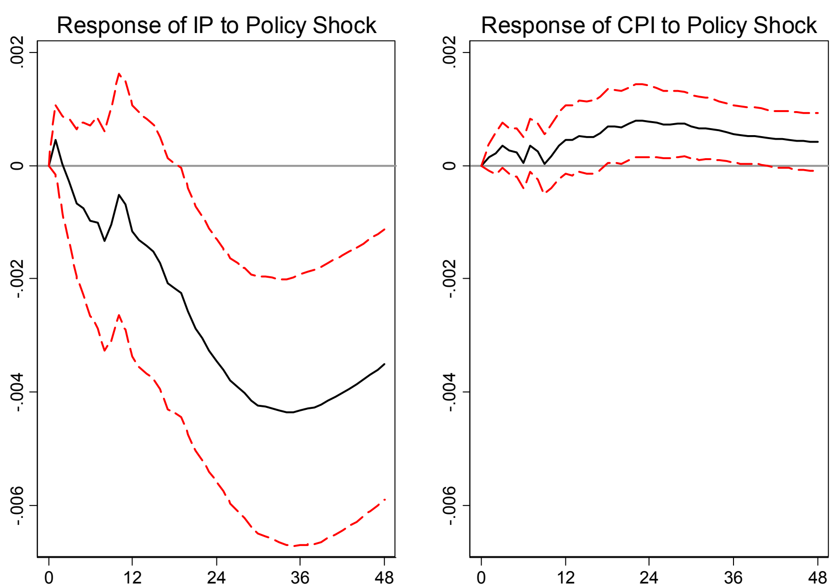
**Figure A3. Modified Lag Structures**
**Panel A.**

Barakchian and Crowe (6 lags). 1988:12-2008:06



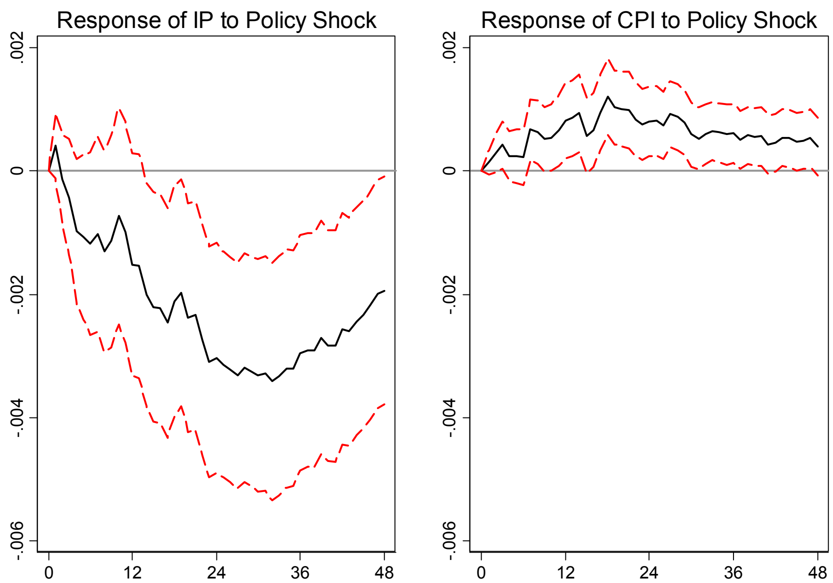

**Panel B.**

Barakchian and Crowe (12 lags). 1988:12-2008:06



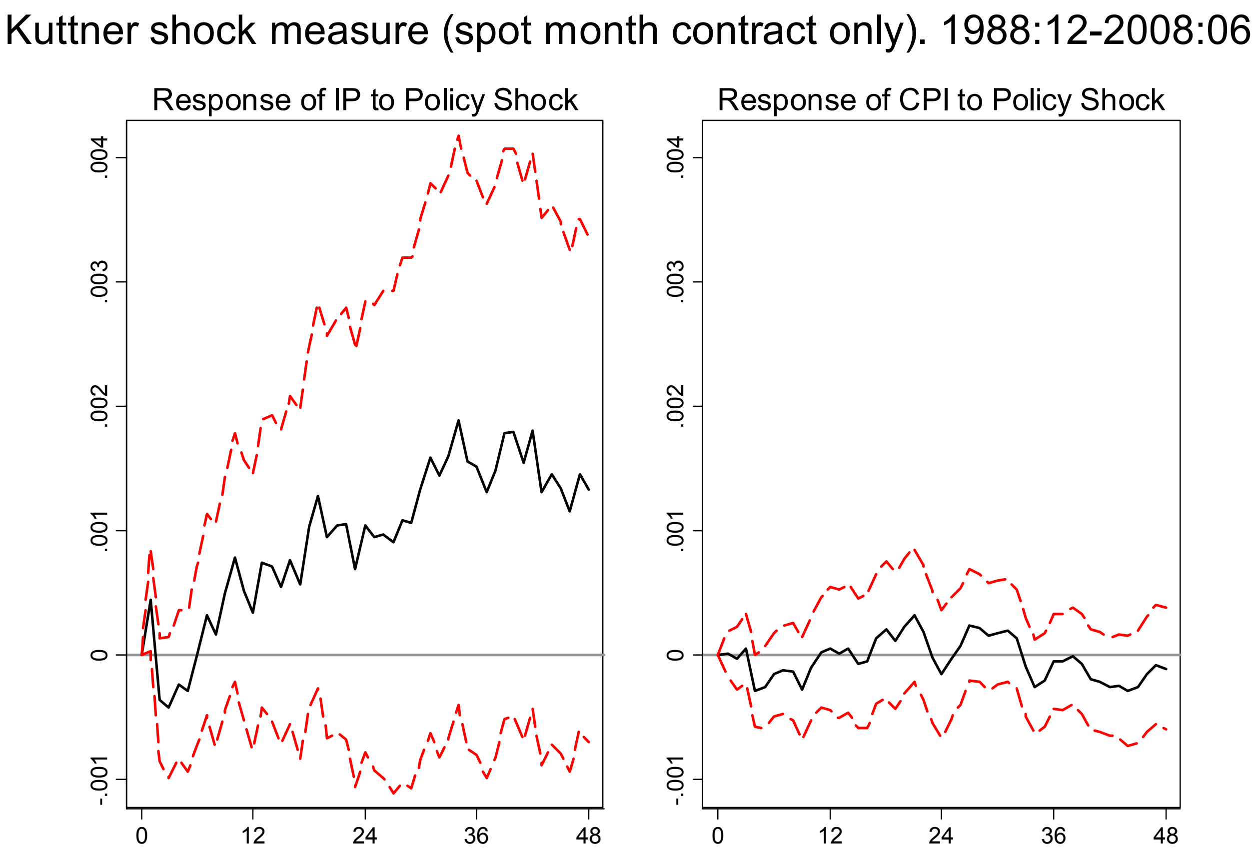
**Panel C.**

Barakchian and Crowe (24 lags). 1988:12-2008:06



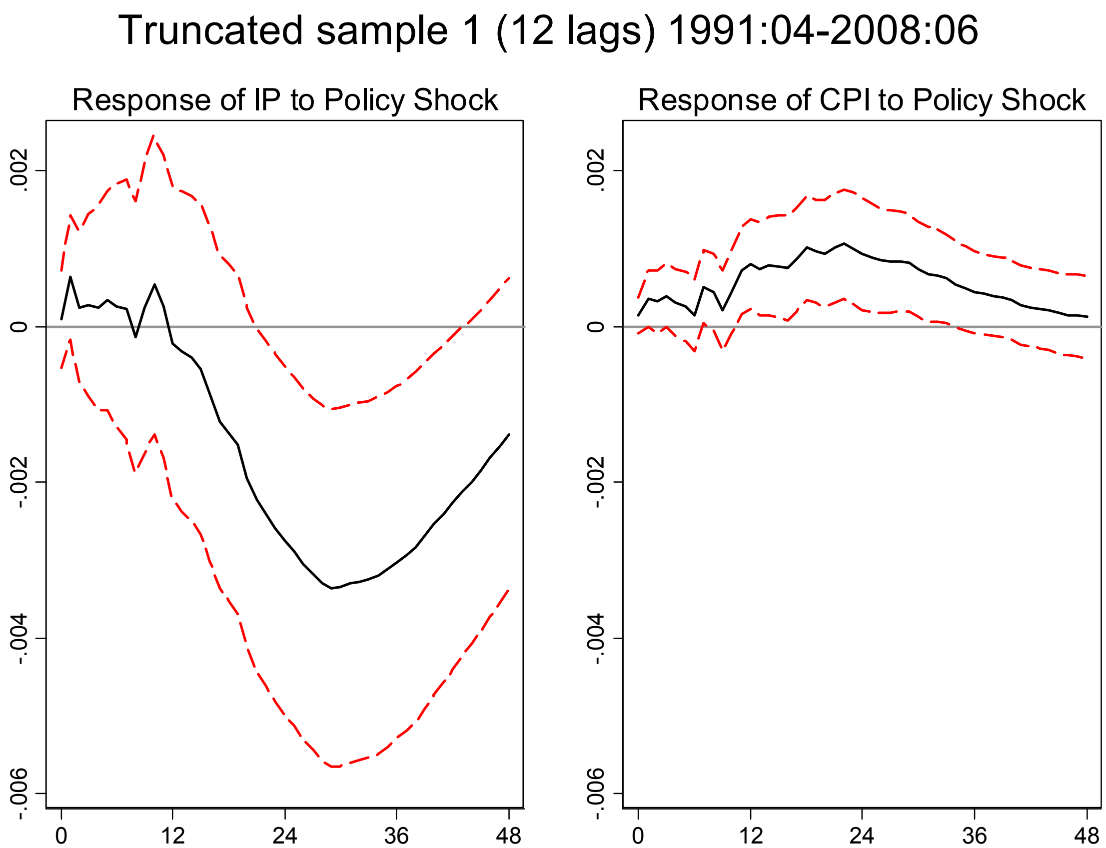

Structural VAR (Monthly data, 3 endogenous variables plus constant and linear time trend, 6, 12 and 24 lags, respectively).
Variables ordered as industrial production, consumer prices, both seasonally adjusted and in logs, and our shock measure, cumulated. Graphs show response of industrial production and CPI to a one standard deviation positive shock to the policy measure.
Structural shocks obtained via Cholesky decomposition.
Two Standard Error bands produced by parametric bootstrapping.

**Figure A4. Using Kuttner (2001) Approach (Spot Month Contract Shock)**



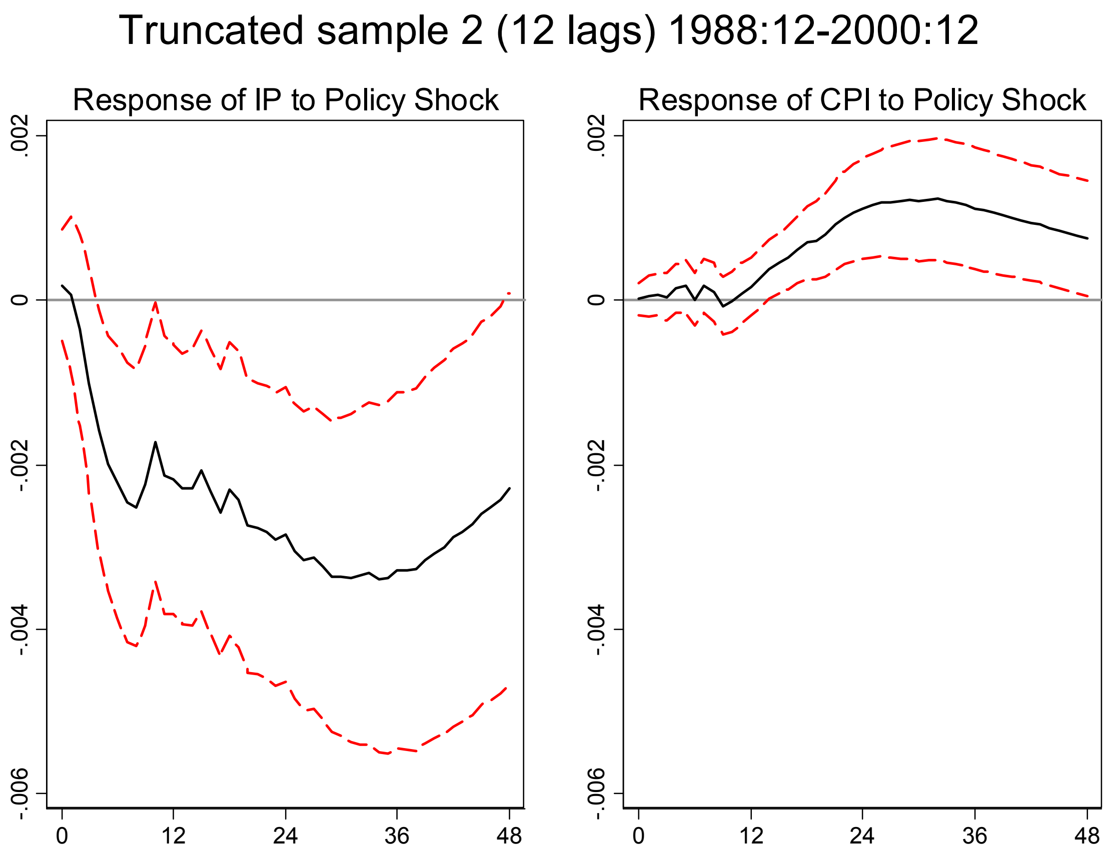

Structural VAR (Monthly data, 3 endogenous variables plus constant and linear time trend, 36 lags).
Variables ordered as industrial production, consumer prices, both seasonally adjusted and in logs, and the Kuttner (2001) shock measure, based on the spot month forward contract, cumulated. Graphs show response of industrial production and CPI to a one standard deviation positive shock to the policy measure.
Structural shocks obtained via Cholesky decomposition.
Two Standard Error bands produced by parametric bootstrapping.

**Figure A5. Truncated Sample (1)**



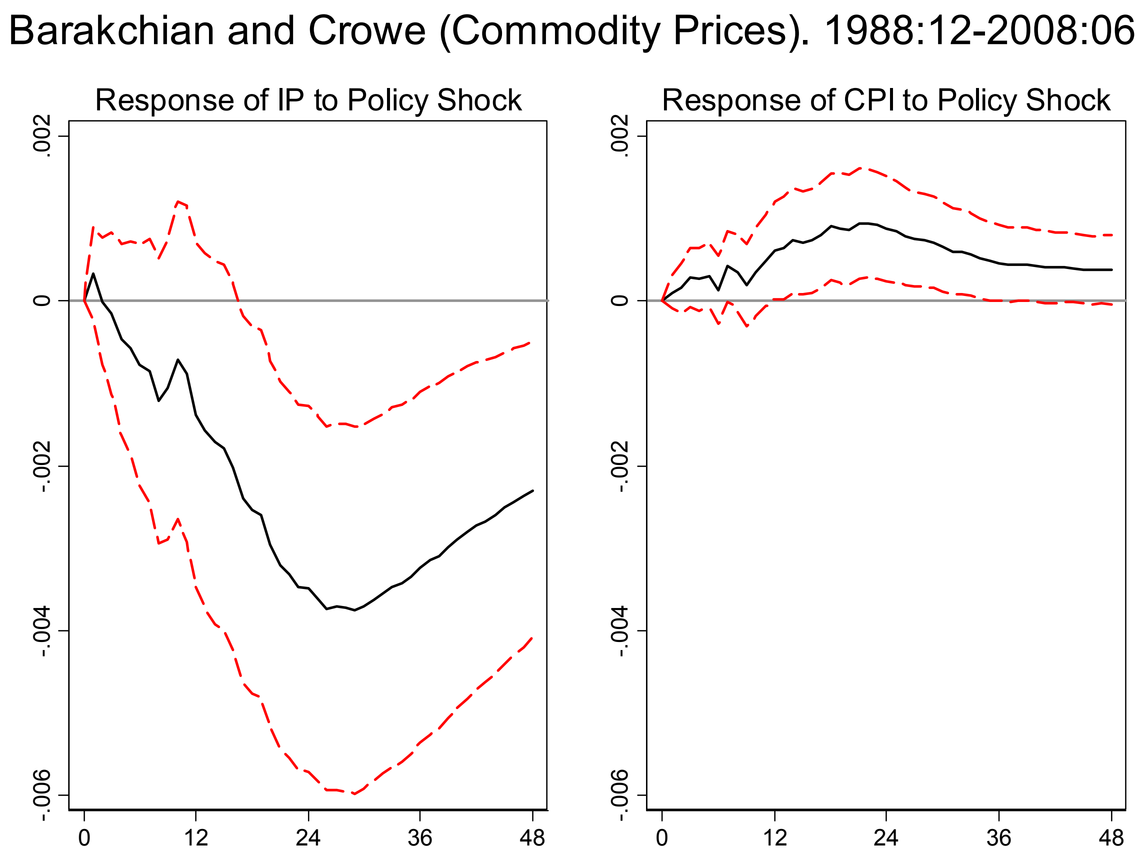

Structural VAR (Monthly data, 3 endogenous variables plus constant and linear time trend, 12 lags).
Sample starts in April 1991, reflecting the NBER dating of the end of the 1990-91 recession in 1991Q1.
Variables ordered as industrial production, consumer prices, both seasonally adjusted and in logs, and our shock measure, cumulated. Graphs show response of industrial production and CPI to a one standard deviation positive shock to the policy measure.
Structural shocks obtained via Cholesky decomposition.
Two Standard Error bands produced by parametric bootstrapping.

**Figure A6. Truncated Sample (2)**



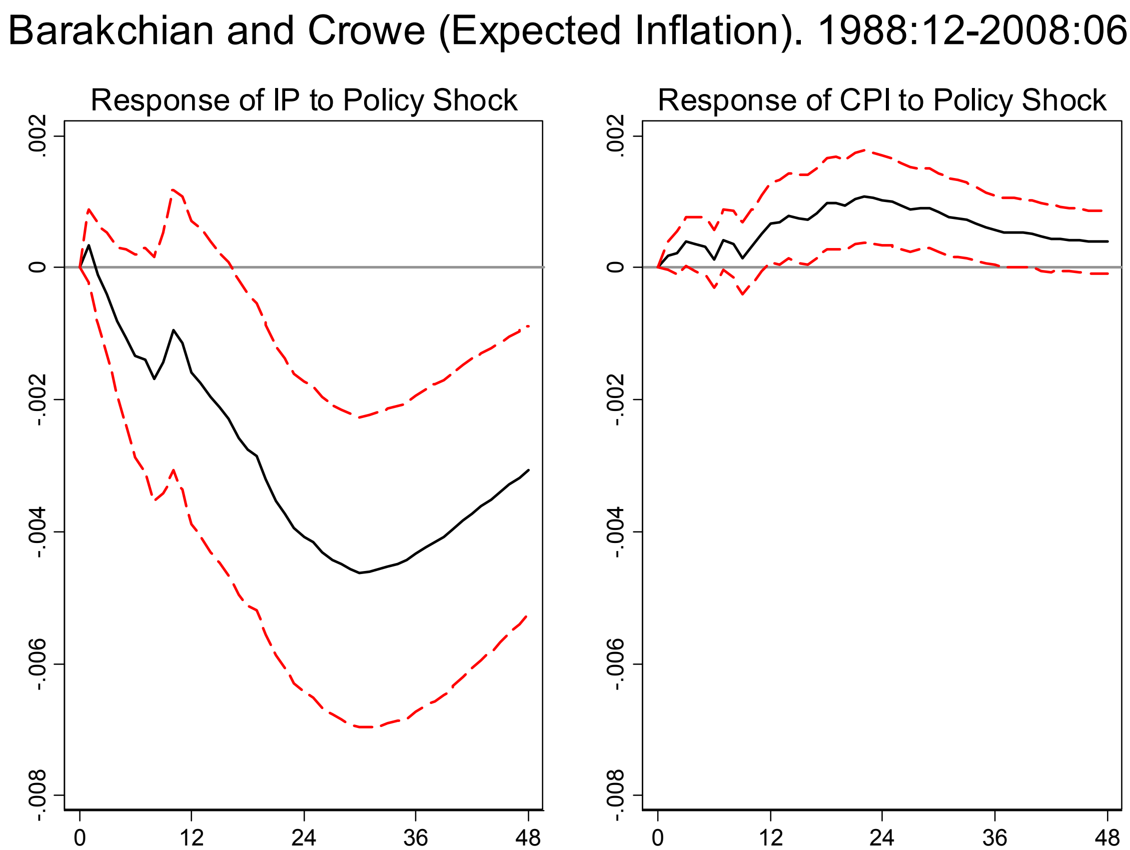

Structural VAR (Monthly data, 3 endogenous variables plus constant and linear time trend, 12 lags).
Sample ends in December 2000, reflecting the NBER dating of the start of the 2001 recession in 2001Q1.
Variables ordered as industrial production, consumer prices, both seasonally adjusted and in logs, and our shock measure, cumulated. Graphs show response of industrial production and CPI to a one standard deviation positive shock to the policy measure.
Structural shocks obtained via Cholesky decomposition.
Two Standard Error bands produced by parametric bootstrapping.

**Figure A7. Commodity Prices**



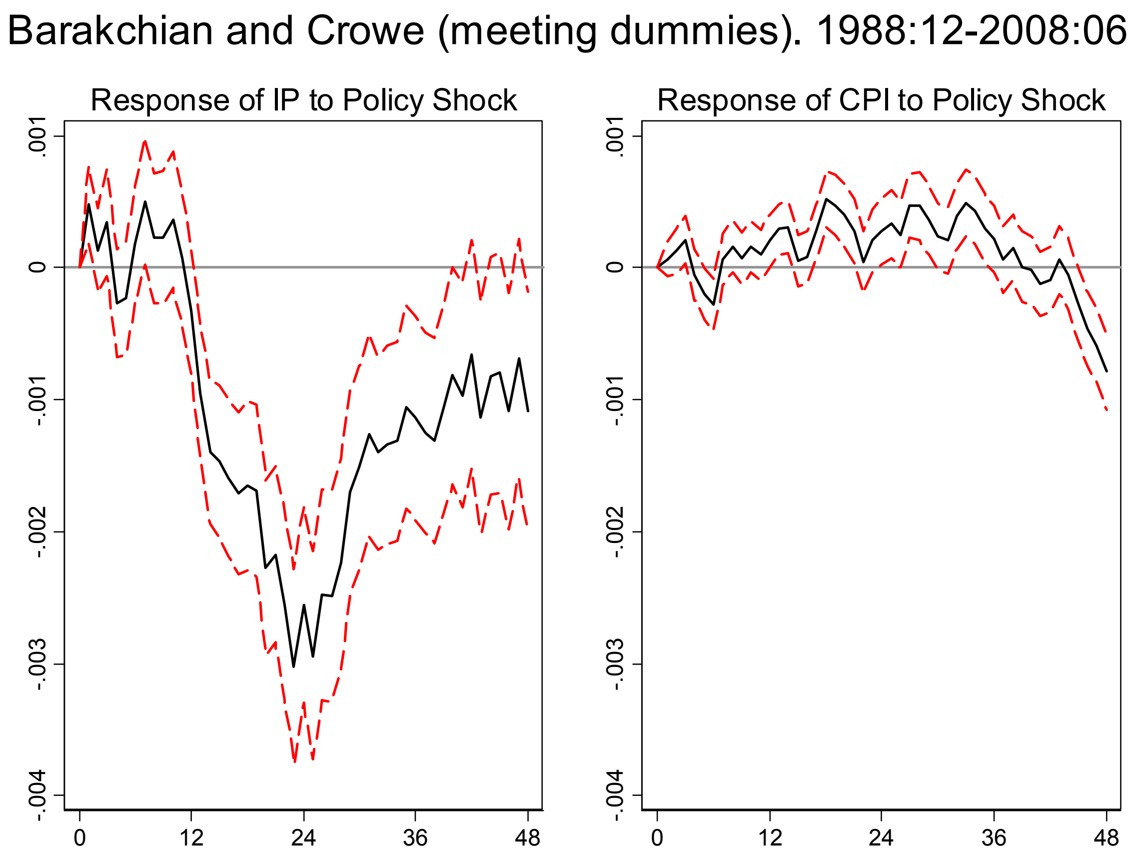

Structural VAR (Monthly data, 4 endogenous variables plus constant and linear time trend, 12 lags).
Variables ordered as commodity prices, industrial production, consumer prices, all seasonally adjusted and in logs, and our shock measure, cumulated. Graphs show response of industrial production and CPI to a one standard deviation positive shock to the policy measure.
Structural shocks obtained via Cholesky decomposition.
Two Standard Error bands produced by parametric bootstrapping.

**Figure A8. Expected Inflation**




Structural VAR (Monthly data, 4 endogenous variables plus constant and linear time trend, 12 lags).
Variables ordered as expected one-quarter ahead inflation, industrial production and consumer prices (both seasonally adjusted and in logs), and our shock measure, cumulated. Graphs show response of industrial production and CPI to a one standard deviation positive shock to the policy measure.
Structural shocks obtained via Cholesky decomposition.
Two Standard Error bands produced by parametric bootstrapping.

**Figure A9. Meeting Date Dummies**

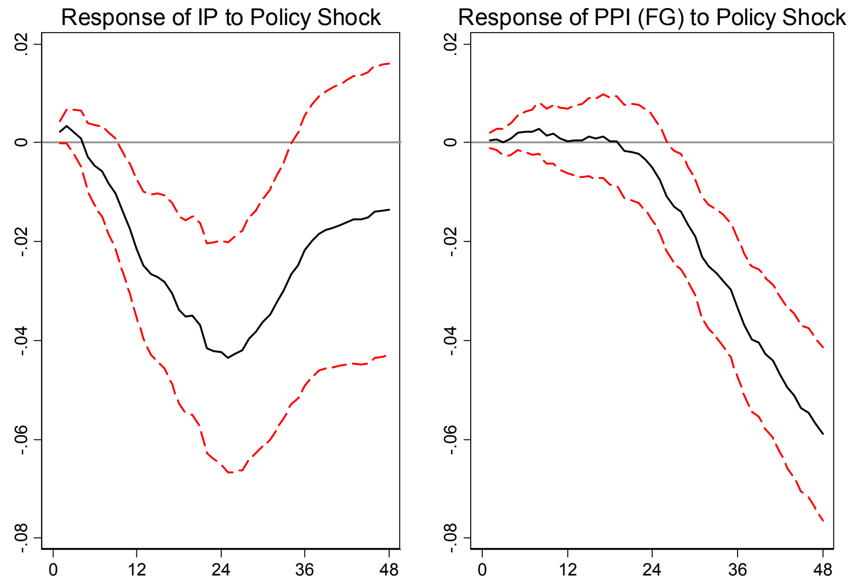



Structural VAR (Monthly data, 3 endogenous variables (36 lags) plus constant, linear time trend and meeting dummies (plus 12 lags of each dummy) as described in the text).
Variables ordered as industrial production, consumer prices (both seasonally adjusted and in logs), and our shock measure, cumulated. Graphs show response of industrial production and CPI to a one standard deviation positive shock to the policy measure.
Structural shocks obtained via Cholesky decomposition.
Two Standard Error bands produced by parametric bootstrapping.

# Figure A10. Romer and Romer Single equation results

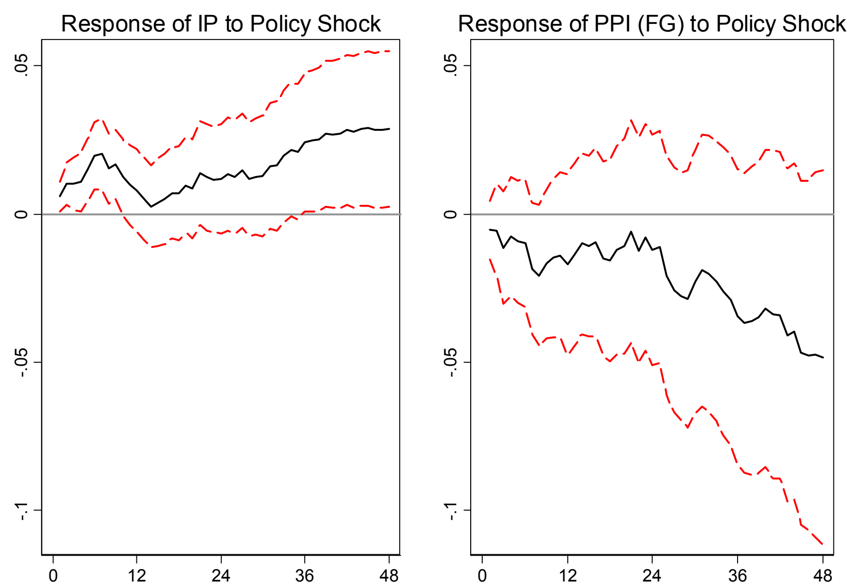
## Panel A.

### Romer and Romer. 1969:01-1996:12



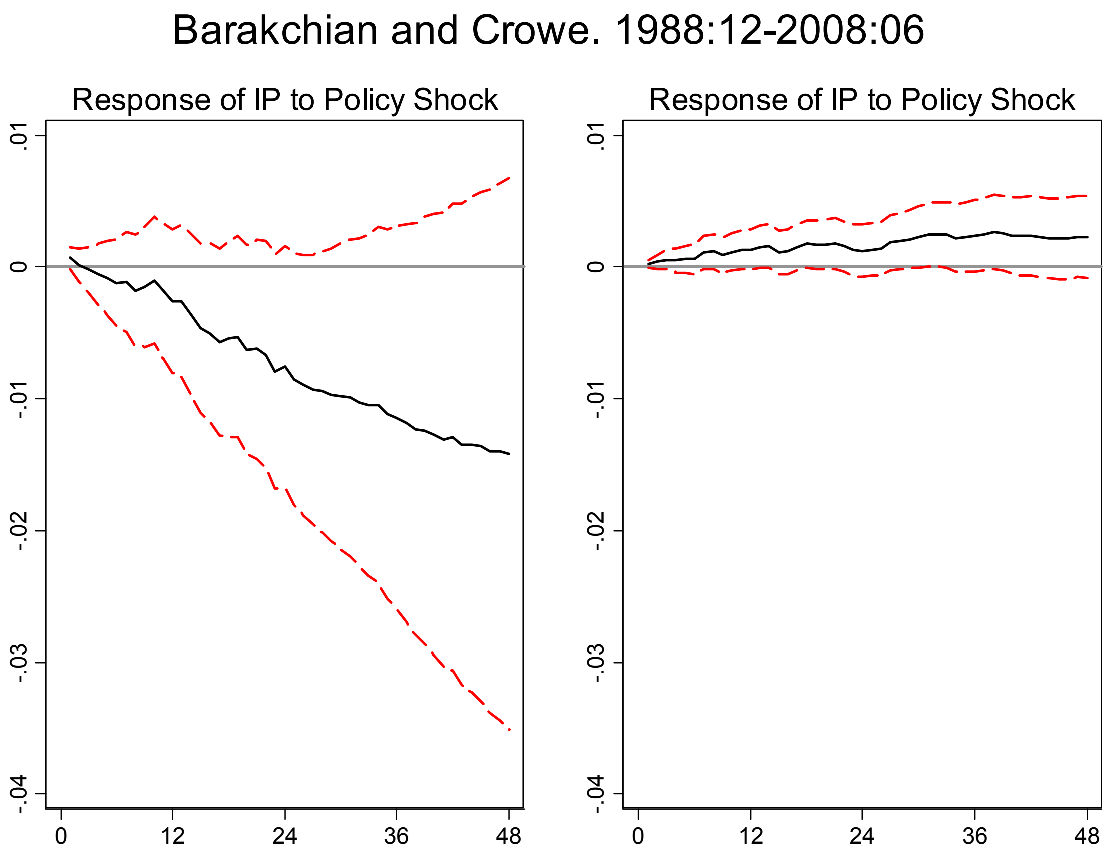

## Panel B.

### Romer and Romer. 1988:12-2008:06




Single Equation estimation (Monthly data, change in log dependent variable regressed on 36 lags of Romer and Romer shock measure (48 lags for price equation) and 24 lags of DV) as described in text.
Graphs show response of industrial production and PPI (finished goods) to a one unit (100 basis points) positive shock to the policy measure (cumulative dynamic multiplier functions).
Two Standard Error bands produced by Monte Carlo methods (500 replications) following the methodology in Romer and Romer (2004, footnote 17).

**Figure A11. Our Shock Measure: Single Equation results**




Single Equation estimation (Monthly data, change in log dependent variable regressed on 36 lags of our shock measure and 24 lags of DV) as described in text.
Graphs show response of industrial production and CPI to a one unit (equal to one standard deviation for non-scaled policy shock obtained from factor model) positive shock to the policy measure (cumulative dynamic multiplier functions).
Two Standard Error bands produced by Monte Carlo methods (500 replications) following the methodology in Romer and Romer (2004, footnote 17).

**References**

Bagliano, F. C. and C.A. Favero, 1998, “Measuring monetary policy with VAR models an evaluation,” *European Economic Review*, Vol. 42, pp. 1113–1140.

Bernanke, Ben S., Jean Boivin and Piotr Eliasz, 2005, “Measuring the effects of monetary policy: A factor-augmented vector autoregression (favar) approach,” *Quarterly Journal of Economics*, Vol. 120, pp. 387–422.

Bernanke, Ben S. and Kenneth N. Kuttner, 2005, “What explains the stock market's reaction to federal reserve policy?” *Journal of Finance,* Vol. 60, pp. 1221–57.

Bernanke, Ben S. and Jean Boivin, 2003, “Monetary policy in a data-rich environment,” *Journal of Monetary Economics* Vol. 50, pp. 525–46.

Bernanke, Ben S. and Mihov, I., 1998, “Measuring monetary policy,” *Quarterly Journal of Economics,* Vol. 113, pp. 869-902.

Blanchard, O.J. and D. Quah, 1989, “The dynamic effects of aggregate supply and demand disturbances,” *American Economic Review* 79, pp. 655-73.

Boivin, Jean and Marc P. Giannoni, 2006, “Has Monetary Policy Become More Effective?” *Review of Economics and Statistics*, Vol. 88, No.3, pp. 445–62.

Castelnuovo, Efrem and Paolo Surico, 2006, “The Price Puzzle: Fact or Artefact?” Bank of England Working Paper No. 288, London: Bank of England.

Christiano, Lawrence J., Martin Eichenbaum and Charles L. Evans, 1996, “The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds,” *Review of Economics and Statistics* Vol. 78, No. 1, pages 16–34.

———, 1999, “Monetary policy shocks: What have we learned and to what end?” In: J.B. Taylor and M. Woodford, Editors, *Handbook of Macroeconomics*, Vol. 1. Amsterdam: Elsevier B.V.

Clarida, R., J. Gali and M. Gertler, 1999, “The science of monetary policy: a new Keynesian perspective,” *Journal of Economic Literature*, Vol. 37, pp. 1661–707.

———, 2000, “Monetary policy rules and macroeconomic stability: evidence and some theory,” *Quarterly Journal of Economics*, Vol. 65, pp. 147–80.

Crowe, Christopher and Ellen E. Meade, 2007, “The Evolution of Central Bank Governance around the World,” *Journal of Economic Perspectives,* Vol. 21, No. 4, pp. 69–90.

Eichenbaum, Martin, 1992, "Comment on Interpreting The Macroeconomic Time Series Facts: The Effects of Monetary Policy," *European Economic Review*, Vol. 36, No. 5, pp. 1001–11.

Faust, J., E. Swanson and J. Wright, 2004, "Identifying VARs based on high-frequency futures data," *Journal of Monetary Economics* 6, pp. 1107–31.

Giordani, Paolo, 2004, "An alternative explanation of the price puzzle," *Journal of Monetary Economics,* Vol. 51, pp. 1271–296.

Greene, William H., 2008, *Econometric Analysis*, Sixth Edition, Upper Saddle River, NJ: Pearson Prentice Hall.

http://www.princeplaza-makati.com/location.php, Gürkaynak, R., Sack, B. and E. Swanson, 2005, "Do Actions Speak Louder than Words? The Response of Asset Prices to Monetary Policy Actions and Statements," *International Journal of Central Banking* Vol. 1, pp. 55–94.

Hamilton, James D., 2008, "Daily monetary policy shocks and new home sales," *Journal of Monetary Economics,* Vol. 55, pp. 1171–90.

Hanson, Michael S., 2004, "The price puzzle reconsidered," *Journal of Monetary Economics,* Vol. 51, pp. 1385–413.

Kuttner, Kenneth N., 2001, "Monetary policy surprises and interest rates: evidence from the federal funds futures market," *Journal of Monetary Economics*, Vol. 47, pp. 523–44.

Orphanides, A., 2003, "Historical monetary policy analysis and the Taylor rule," *Journal of Monetary Economics*, Vol. 50, pp. 983–1022.

Ozlale, Umit, 2003, "Price stability vs. output stability: tales of federal reserve administrations," *Journal of Economic Dynamics and Control* 27, pp. 1595-610.

Piazzesi, Monika, 2010, "Affine term structure models," In *Handbook of financial econometrics*, ed. L. P. Hansen and Y. Ait-Sahalia, Amsterdam: North-Holland, pp. 691–766.

Piazzesi, Monika and Eric T. Swanson, 2008, "Futures prices as risk-adjusted forecasts of monetary policy," *Journal of Monetary Economics,* Vol. 55, pp. 677–91.

Romer, Christina D. and David H. Romer, 1989, "Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz," in Olivier J. Blanchard and Stanley Fischer, eds., *NBER Macroeconomics Annual* 1989, Cambridge, MA: MIT Press, pp. 121–70.

Romer, Christina D. and David H. Romer, 2000, "Federal Reserve Information and the Behavior of Interest Rates," *American Economic Review*, Vol. 90, No. 3, pp. 429–57.

Romer, Christina D. and David H. Romer, 2004, "A new measure of monetary shocks: derivation and implications," *American Economic Review,* Vol. 94, pp. 1055–84.

Rudebusch, G. D., 1998, "Do measures of monetary policy in a VAR make sense?" *International Economic Review*, Vol. 39, pp. 907–31.

Sims, Christopher A., 1980, "Macroeconomics and Reality," *Econometrica*, Vol. 48, No 1, pp. 1-48.

———, 1992, "Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy," *European Economic Review*, Vol. 36, No. 5, pp. 975-1000.

Sims, Christopher A. and Tao Zha, 2006, "Does monetary policy generate recessions?" *Macroeconomic Dynamics*, Vol. 10, pp. 231–72.

Söderström, U., 2001, Predicting monetary policy with Federal funds futures prices, *Journal of Futures Markets* 214, pp. 377-91.

Taylor John, B., 1993, Discretion versus policy rules in practice. In: *Carnegie-Rochester Conference Series on Public Policy*, Vol. 39, pp. 195–214.

Thapar, Aditi, 2008, Using private forecasts to estimate the effects of monetary policy. *Journal of Monetary Economics*, Vol. 55, pp. 806–24.

Uhlig, H., 2005, "What are the effects of monetary policy on output? Results from an agnostic identification procedure," *Journal of Monetary Economics*, Vol. 52, pp. 381–419.